



UNIVERZITET U NOVOM SADU
FAKULTET TEHNIČKIH NAUKA U
NOVOM SADU



Olivera Lazić

**Primena veštačke inteligencije u optimizaciji
uzgoja suncokreta: Detekcija stresa usled suše i
predviđanje visine prinosa ulja**

DOKTORSKA DISERTACIJA

Novi Sad, 2026.

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА¹

Врста рада:	Докторска дисертација
Име и презиме аутора:	Оливера Лазић
Ментор (титула, име, презиме, звање, институција):	Др Александар Купусинац, редовни професор, Факултет техничких наука, Универзитет у Новом Саду
Наслов рада:	Примена вештачке интелигенције у оптимизацији узгоја сунцокрета: Детекција стреса услед суше и предвиђање висине приноса уља
Језик и писмо рада:	Српски (латиница)
Физички опис рада:	Унети број: Страница: 122 Поглавља: 7 Референци: 212 Табела: 16 Слика: 36 Графикона: 0 Прилога: 1
Научна област:	Електротехничко и рачунарско инжењерство
Ужа научна област (научна дисциплина):	Примењене рачунарске науке и информатика
Кључне речи / предметна одредница:	Вештачка интелигенција, дубоко учење, детекција стреса услед суше, машинско учење, предвиђање висине приноса уља, прецизна пољопривреда, слике из ризотрона, сунцокрет, трансфер учење.
Апстракт на језику рада:	<p>Са порастом глобалне људске популације расте и потражња за храном и водом, што врши притисак на пољопривредни сектор да повећа продуктивност уз минимизацију негативног утицаја на животну средину. Дисертација се бави применом технологија вештачке интелигенције у оптимизацији производње сунцокрета кроз два кључна изазова: детекцију стреса услед суше код биљака сунцокрета и предвиђање висине приноса сунцокретовог уља.</p> <p>За детекцију стреса услед суше, спроведени су експерименти са 12 инбрид линија сунцокрета одабраних на основу њихове толеранције на сушу, продуктивности и отпорности на болести. Биљке су узгајане у контролисаним условима у ризотронима, омогућавајући неинвазивно праћење одговора корена на различите нивое влаге у земљишту (70%, 50% и 42% гравиметријског садржаја воде). Развијени су модели засновани на дубоком и трансфер учењу како би се открио сушни стрес на основу слика делова сунцокрета. Упоредене су различите конфигурације модела како би се добио најпрецизнији, најпоузданији и најефикаснији модел.</p> <p>За предвиђање приноса уља, спроведени су теренски огледи на четири локације са различитим агроколошким условима у панонском региону</p>

¹ Аутор докторске дисертације потписао је и приложио следеће Обрасце:

5б – Изјава о ауторству;

5в – Изјава о истовестности штампане и електронске верзије докторског рада и дозвола за објављивање личних података;

5г – Изјава о коришћењу.

Ове Изјаве се чувају у институцији у штампаном и електронском облику и не коришће се са радом.

	<p>Србије. Коришћени су подаци за 1250 хибрида, при чему је за сваки хибрид спроведено по три понављања на све четири локације. Подаци о приносу семена, садржају уља и факторима животне средине као што су температура, падавине и својства земљишта, прикупљени су током две вегетационе сезоне (2018. и 2019. године). Обучени су различити модели машинског и дубоког учења, а њихови резултати су евалуирани како би се пронашао модел са највећом тачношћу.</p> <p>Основни доприноси представљају увођење новог приступа заснованог на вештачкој интелигенцији, за откривање показатеља стреса услед суше код биљака сунцокрета, развијање предиктивног модела за висину приноса сунцокретовог уља на основу различитих еколошких и агрономских фактора и креирање свеобухватних скупова података слика корена и изданака сунцокрета, као и теренских података о хибридима сунцокрета у току две вегетационе сезоне.</p>
Датум прихватања теме од стране надлежног већа:	26.02.2026.
Датум одбране: (Попуњава накнадно институција)	
Чланови комисије: (титула, име, презиме, звање, институција)	<p>Председник: др Срђан Попов, редовни професор, Универзитет у Новом Саду, Факултет техничких наука</p> <p>Члан: др Дуња Врбашки, доцент, Универзитет у Новом Саду, Факултет техничких наука</p> <p>Члан: др Ксенија Дорословачки, редовни професор, Универзитет у Новом Саду, Факултет техничких наука</p> <p>Члан: др Јелена Јоцковић, научни сарадник, Институт за ратарство и повртарство, Нови Сад</p> <p>Ментор: др Александар Купусинац, редовни професор, Универзитет у Новом Саду, Факултет техничких наука</p>
Напомена:	

KEY WORD DOCUMENTATION²

Document type:	Doctoral dissertation
Author:	Olivera Lazić
Supervisor (title, first name, last name, position, institution)	PhD, Aleksandar Kupusinac, full professor, Faculty of Technical Sciences, University of Novi Sad
Thesis title in English:	Artificial Intelligence in Optimizing Sunflower Cultivation: Drought Stress Detection and Oil Yield Prediction
Language and script:	Serbian (Latin)
Physical description:	Number of: Pages: 122 Chapters: 7 References: 212 Tables: 16 Illustrations: 36 Graphs: 0 Appendices: 1
Scientific field:	Electrical and computer engineering
Scientific subfield (scientific discipline):	Applied Computer Science and Informatics
Subject, Key words:	Artificial intelligence, deep learning, drought stress detection, machine learning, oil yield prediction, precision agriculture, rhizotron imaging, sunflower, transfer learning.
Abstract in English:	<p>As the world population continues to grow, so does the demand for food and water, placing increasing pressure on the agricultural sector to enhance productivity while minimizing environmental impacts.</p> <p>The dissertation focuses on applying artificial intelligence technologies to optimize sunflower production by addressing two critical challenges: sunflower drought stress detection and sunflower oil yield prediction.</p> <p>To detect drought stress, experiments were conducted on 12 sunflower inbred lines selected for their drought tolerance, productivity, and disease resistance. The plants were cultivated under controlled conditions in rhizotrons, enabling non-invasive monitoring of root system responses to varying soil moisture levels (70%, 50%, and 42% gravimetric water content). Deep and transfer learning-based models were developed to detect drought stress using images of sunflower. Various model configurations were tested and fine-tuned to achieve the most accurate, reliable, and efficient model for stress detection.</p> <p>To predict oil yield, field trials were conducted at four locations with diverse agro-ecological conditions in the Pannonian region of Serbia. Data from 1250 sunflower hybrids was collected, with three replicates for each hybrid across all locations. The dataset included measurements of seed yield, oil content, and environmental factors such as temperature, precipitation, and soil</p>

² The author of the doctoral dissertation has signed the following Statements:

56 – Statement on the authorship,

5B – Statement that the printed and e-version of the doctoral dissertation are identical and authorization to use personal data,

5r – Copyright statement.

The paper and e-versions of Statements are held at the institution and are not included into the printed thesis.

	<p>properties, recorded over two growing seasons (2018 and 2019). Different machine and deep learning models were trained, and their performance was evaluated to identify the model with the highest accuracy.</p> <p>The main contributions of the dissertation include the development of a novel artificial intelligence-based approach for the detection of drought stress indicators in sunflower plants, the development of a predictive model for estimating sunflower oil yield based on various environmental and agronomic factors, and the creation of comprehensive datasets comprising sunflower root and shoot images, as well as field data on sunflower hybrids collected over two growing seasons.</p>
Date of endorsement by the scientific board:	26.02.2026.
Date of defence: (Filled in by the institution)	
Thesis defence board: (title, first name, last name, position, institution)	<p>Chair: Srđan Popov, PhD, full professor, University of Novi Sad, Faculty of Technical Sciences</p> <p>Member: Dunja Vrbaški, PhD, assistant professor, University of Novi Sad, Faculty of Technical Sciences</p> <p>Member: Ksenija Doroslovački, PhD, full professor, University of Novi Sad, Faculty of Technical Sciences</p> <p>Member: Jelena Jocković, PhD, research associate, Institute of Field and Vegetable Crops, Novi Sad</p> <p>Mentor: Aleksandar Kupusinac, PhD, full professor, University of Novi Sad, Faculty of Technical Sciences</p>
Note:	

Bošku i Teodoru ♥

Sadržaj

1. Uvod	1
1.1 Motivacija	1
1.2 Problem istraživanja	3
1.3 Predmet istraživanja	4
1.3.1 Predmet istraživanja za detekciju stresa usled suše	5
1.3.2 Predmet istraživanja za predviđanje visine prinosa ulja	5
1.4 Ciljevi istraživanja	6
1.4.1 Cilj istraživanja za detekciju stresa usled suše	6
1.4.2 Cilj istraživanja za predviđanje visine prinosa ulja	6
1.5 Hipoteze u istraživanju	7
1.5.1 Hipoteze za detekciju stresa usled suše	7
1.5.2 Hipoteze za predviđanje visine prinosa ulja	7
1.6 Struktura teze	7
2. Pregled literature i aktuelno stanje u oblasti	9
2.1 Primena veštačke inteligencije u agronomiji	9
2.2 Primena veštačke inteligencije u detekciji stresa usled suše	11
2.3 Primena veštačke inteligencije u predviđanju visine prinosa ulja	15
2.4 Problemi i njihova potencijalna rešenja u primeni veštačke inteligencije u agronomiji	19
3. Teorijske osnove	21
3.1 Uvod u veštačku inteligenciju	21
3.1.1 Definicija i osnovna podela veštačke inteligencije	21
3.1.2 Istorijski razvoj veštačke inteligencije	22
3.1.3 Značaj i primenljivost veštačke inteligencije	24
3.2 Mašinsko učenje	26
3.2.1 Nadgledano učenje	27
3.2.1.1 Metod potpornih vektora za regresiju	29
3.2.1.2 Metod K-najbližih suseda	30
3.2.1.3 Veštačka neuronska mreža	30
3.2.1.4 Slučajna šuma	30
3.2.2 Nenadgledano učenje	31
3.2.3 Polu-nadgledano učenje	31
3.2.4 Učenje uz podsticaje	32
3.3 Veštačke neuronske mreže	33
3.4 Duboko učenje	35
3.4.1 Uvod u duboko učenje	35
3.4.2 Razlike između dubokog i mašinskog učenja	37
3.5 Konvolucione neuronske mreže	37
3.5.1 Arhitektura konvolucione neuronske mreže	38
3.5.2 Najčešće korišćene vrste konvolucionih neuronskih mreža ..	41
3.6 Transferno učenje	43

3.7 Tehnike pripreme podataka	44
3.7.1 Eksploratorna analiza podataka	44
3.7.2 Tehnike augmentacije podataka	44
4. Metodologija	46
4.1 Metodologija za detekciju stresa usled suše	46
4.1.1 Opis eksperimentalne postavke	46
4.1.2 Opis skupa podataka	48
4.1.3 Obrada i obogaćivanje podataka	50
4.1.3.1 Klasna neuravnoteženost i prilagođavanje eksperimentalnog dizajna	51
4.1.3.2 Generisanje dodatnih uzoraka deljenjem slika korenovog sistema	52
4.1.4 Odabir modela dubokog učenja	52
4.1.5 Priprema i obuka modela dubokog učenja	54
4.1.6 Metode evaluacije	56
4.2 Metodologija za predviđanje visine prinosa ulja	57
4.2.1 Opis eksperimentalne postavke	58
4.2.2 Opis skupa podataka	59
4.2.3 Eksploratorna analiza skupa podataka	61
4.2.4 Korišćeni modeli mašinskog učenja	65
4.2.5 Obuka modela mašinskog učenja	66
4.2.6 Metode evaluacije	67
4.3 Implementacija softverskog rešenja	68
5. Rezultati	70
5.1. Rezultati detekcije stresa usled suše	70
5.1.1 Rezultati eksperimenta u inicijalnoj postavci	70
5.1.1.1 Poređenje performansi modela i uticaja formata slika	70
5.1.1.2 Vremenska efikasnost modela	72
5.1.1.3 Vizuelna analiza obuke modela	72
5.1.1.4 Analiza performansi i izbor optimalnog modela	75
5.1.2 Evaluacija performansi MobileNet modela uz različite pristupe augmentacije podataka	75
5.1.2.1 Rezultati za slike u boji	76
5.1.2.2 Rezultati za slike u sivim tonovima	76
5.1.3 Vizuelna analiza rezultata MobileNet modela	77
5.2. Rezultati predikcije visine prinosa ulja	81
5.2.1. Performanse regresionih algoritama	81
5.2.2. Vizuelna analiza predikcija	83
5.2.3. Analiza važnosti karakteristika	84
5.2.4 Statistička analiza i uticaj karakteristika na prinos ulja	86
6. Diskusija	92
6.1 Diskusija rezultata za detekciju stresa usled suše	92
6.2 Diskusija rezultata za predviđanje visine prinosa ulja	94

6.2.1 Analiza karakteristika hibrida sa najvećim prinosom ulja	94
6.2.2 Diskusija primene metoda mašinskog učenja za predviđanje prinosa ulja	95
7. Zaključak	97
Dodatak	99
Reference	101

Lista skraćenica

AI – Artificial Intelligence – veštačka inteligencija

ANN – Artificial Neural Network – veštačka neuronska mreža

AUC – Area Under the Curve – površina ispod krive

CNN – Convolutional Neural Network – konvoluciona neuronska mreža

DL – Deep Learning – duboko učenje

DNN – Deep Neural Network – duboka neuronska mreža

DT – Decision Tree – stablo odlučivanja

EDA – Exploratory Data Analysis – eksploratorna analiza podataka

GAM – Generalized Additive Model – generalizovani aditivni model

GAN – Generative Adversarial Network – generativna adverzijalna mreža

GWC – Gravimetric Water Content – gravimetrijski sadržaj vode

ILSVRC – ImageNet Large Scale Visual Recognition Challenge

K-NN – K-Nearest Neighbors – metoda najbližih suseda

LLM – Large Language Model – veliki jezički model

LSTM – Long Short-Term Memory – mreža sa dugoročnom i kratkoročnom memorijom

MAE – Mean Absolute Error – srednja apsolutna greška

ML – Machine Learning – mašinsko učenje

MSE – Mean Squared Error – srednja kvadratna greška

NDVI – Normalized Difference Vegetation Index – normalizovani diferencijalni vegetacioni indeks

NMR – Nuclear Magnetic Resonance – nuklearna magnetna rezonanca

RF – Random Forest – slučajna šuma

RFR – Random Forest Regression – regresija metodom slučajne šume

RMSE – Root Mean Squared Error – koren srednje kvadratne greške

RNN – Recurrent Neural Network – rekurentna neuronska mreža

ROC – Receiver Operating Characteristic – karakteristika operativnosti prijemnika

SLR – Stepwise Linear Regression – postupna linearna regresija

SVM – Support Vector Machine – mašina potpornih vektora

SVR – Support Vector Regression – regresija pomoću potpornih vektora

TL – Transfer Learning – transferno učenje

UAV – Unmanned Aerial Vehicle – bespilotna letelica

Indeks slika

Slika 1. Grafički prikaz Turingovog testa.	23
Slika 2. Međusobni odnos veštačke inteligencije, mašinskog učenja i dubokog učenja.	26
Slika 3. Biološki neuron i matematički model veštačkog neurona (Ahmed i Lindblad, 2022).	34
Slika 4. Obuka višeslojne neuronske mreže (Rao et al., 2015).	35
Slika 5. Arhitektura konvolucione neuronske mreže (Singh, 2024).	39
Slika 6. Konvolucija u konvolucionim neuronskim mrežama (Nagib et al., 2022).	39
Slika 7. Eksperimentalna postavka - biljke suncokreta u klima komori.	47
Slika 8. Biljka suncokreta u rizotronu.	48
Slika 9. Reprezentativni primeri skupa podataka: (a) fotografija korena suncokreta iz rizotrona; (b) skenirani korenov sistem; (c) skenirani nadzemni deo biljke.	50
Slika 10. Distribucija prinosa ulja za skup podataka iz 2018. godine: (a) prvi podskup podataka; (b) drugi podskup podataka.	62
Slika 11. Distribucija prinosa ulja za skup podataka iz 2019. godine: (a) prvi podskup podataka; (b) drugi podskup podataka.	62
Slika 12. Koeficijenti Pirsonovih korelacija numeričkih varijabli za podatke iz 2018. godine.	63
Slika 13. Koeficijenti Pirsonovih korelacija numeričkih varijabli za podatke iz 2019. godine.	63
Slika 14. Vrednosti visine prinosa ulja u odnosu na lokalitet u 2018. godini.	64
Slika 15. Vrednosti visine prinosa ulja u odnosu na lokalitet u 2019. godini.	65
Slika 16. Rezultati tačnosti i gubitka za deo skupa podataka za obuku u boji.	73
Slika 17. Rezultati tačnosti i gubitka za deo skupa podataka za validaciju u boji.	73
Slika 18. Rezultati tačnosti i gubitka za deo skupa podataka za obuku u sivim tonovima.	74
Slika 19. Rezultati tačnosti i gubitka za deo skupa podataka za validaciju u sivim tonovima.	74
Slika 20. Krive tačnosti i funkcije gubitka MobileNet modela tokom obuke za originalni, augmentovani i ciljano augmentovani skup podataka.	77
Slika 21. Krive tačnosti i funkcije gubitka MobileNet modela tokom validacije za originalni, augmentovani i ciljano augmentovani skup podataka.	78
Slika 22. Matrice konfuzije koje prikazuju performanse različitih MobileNet konfiguracija na test skupu podataka.	78

Slika 23. Kriva operativnih karakteristika prijemnika za različite konfiguracije MobileNet modela.	80
Slika 24. Izmerene i predviđene vrednosti za RFR model za prvi podskup podataka u 2018. (a) i 2019. godini (b).	83
Slika 25. Izmerene i predviđene vrednosti za RFR model za drugi podskup podataka u 2018. (a) i 2019. godini (b).	84
Slika 26. Najrelevantnije karakteristike za RFR u 2018. godini za prvi podskup podataka.	85
Slika 27. Najrelevantnije karakteristike za RFR u 2019. godini za prvi podskup podataka.	85
Slika 28. Najrelevantnije karakteristike za RFR u 2018. godini za drugi podskup podataka.	86
Slika 29. Najrelevantnije karakteristike za RFR u 2019. godini za drugi podskup podataka.	86
Slika 30. Vrednosti visine prinosa ulja u odnosu na karakteristiku zrelost hibrida u 2018. godini.	87
Slika 31. Vrednosti visine prinosa ulja u odnosu na karakteristiku zrelost hibrida u 2019. godini.	88
Slika 32. Vrednosti visine prinosa ulja u odnosu na otpornost na volovod u 2018. godini.	88
Slika 33. Vrednosti visine prinosa ulja u odnosu na otpornost na volovod u 2019. godini.	89
Slika 34. Vrednosti visine prinosa ulja u odnosu na otpornost na plamenjaču u 2018. godini.	90
Slika 35. Vrednosti visine prinosa ulja u odnosu na otpornost na plamenjaču u 2019. godini.	90

Indeks tabela

Tabela 1. Pregled relevantnih radova za detekciju biotičkog stresa.....	13
Tabela 2. Pregled relevantnih radova za detekciju stresa usled suše.....	13
Tabela 3. Pregled relevantnih radova za predviđanje prinosa.....	16
Tabela 4. Pregled korišćenih algoritama mašinskog učenja sa njihovim osnovnim prednostima i ograničenjima.....	29
Tabela 5. Prednosti često korišćenih arhitektura u kontekstu malih skupova podataka.....	42
Tabela 6. Performanse CNN arhitektura na <i>ImageNet</i> skupu (Keras, 2024)..	53
Tabela 7. Optimalan broj epoha i <i>learning rate</i> za korišćene modele.....	56
Tabela 8. Opis prvog podskupa podataka.....	60
Tabela 9. Opis drugog podskupa podataka.....	61
Tabela 10. Poređenje performansi pet odabranih CNN modela (najbolje vrednosti su podebljane).....	70
Tabela 11. Poređenje performansi MobileNet modela treniranog na tri varijante skupa podataka (najbolje vrednosti su podebljane).....	75
Tabela 12. Performanse regresionih modela za prvi podskup podataka.....	82
Tabela 13. Performanse regresionih modela za drugi podskup podataka.....	82
Tabela 14. Hibridi sa najvećim prinosom ulja za 2018. godinu.....	91
Tabela 15. Hibridi sa najvećim prinosom ulja za 2019. godinu.....	91

Rezime

Sa porastom globalne ljudske populacije raste i potražnja za hranom i vodom, što vrši pritisak na poljoprivredni sektor da poveća produktivnost uz minimizaciju negativnog uticaja na životnu sredinu. Klimatske promene su jedan od ključnih faktora u ovom procesu, jer nepredvidivi vremenski obrasci i produžene suše ugrožavaju prinose useva i sigurnost hrane. U tom kontekstu, postoji potreba za razvojem automatizovanih metoda koje bi omogućile agronomima donošenje efikasnih odluka, čak i u nepovoljnim klimatskim uslovima, kao što je nedostatak vode u kombinaciji sa visokim temperaturama. Razvojem modela zasnovanih na veštačkoj inteligenciji moguće je ublažiti navedene izazove kroz optimizovano upravljanje resursima u uslovima ograničene dostupnosti vode i povišenih temperatura.

Ovo istraživanje se fokusira na primenu tehnologija veštačke inteligencije u optimizaciji proizvodnje suncokreta kroz dva ključna izazova: detekciju stresa usled suše kod biljaka suncokreta i predikciju visine prinosa suncokretovog ulja. Suncokret je kao veoma važna uljana biljna vrsta posebno osetljiv na nedostatak vode što utiče na njegov rast i proizvodnju ulja. Tradicionalne metode za utvrđivanje stresa usled suše kod biljaka suncokreta uključuju kombinaciju fizioloških, tehnoloških i molekularnih metoda. Osim toga, tradicionalne metode za procenu visine prinosa ulja za nove hibride suncokreta uglavnom se oslanjaju na poljske ogledе, hemijske analize i statističke metode. Iako pružaju pouzdane rezultate, ove metode su često dugotrajne, skupe i zahtevaju intenzivan ljudski rad. Da bi se prevazišli navedeni problemi, razvijena je metodologija koja kombinuje napredne tehnike mašinskog i dubokog učenja sa eksperimentalnim i terenskim podacima.

Za detekciju stresa usled suše, sprovedeni su eksperimenti za 12 inbred linija suncokreta odabranih na osnovu njihove tolerancije na sušu, produktivnosti i otpornosti na bolesti. Biljke su uzgajane u kontrolisanim uslovima u rizotronima, omogućavajući neinvazivno praćenje odgovora korena na različite nivoe vlage u zemljištu (70%, 50% i 42% gravimetrijskog sadržaja vode). Za snimanje i analizu morfoloških promena u korenovom sistemu i nadzemnim izdancima korišćena je kamera visoke rezolucije i specijalizovani skener. Razvijeni su modeli zasnovani na dubokom i transfernom učenju kako bi se otkrio sušni stres na osnovu slika delova suncokreta. Upoređene su različite konfiguracije modela kako bi se dobio najprecizniji, najpouzdaniji i najefikasniji model.

Za predviđanje prinosa ulja, sprovedeni su terenski ogledi na četiri lokacije sa različitim agroekološkim uslovima u panonskom regionu Srbije. Korišćeni su

podaci za 1250 hibrida, pri čemu je za svaki hibrid sprovedeno po tri ponavljanja na sve četiri lokacije. Podaci o prinosu semena, sadržaju ulja i faktorima životne sredine kao što su temperatura, padavine, svojstva zemljišta prikupljeni su tokom dve vegetacione sezone (2018. i 2019. godine). Obučeni su različiti modeli mašinskog i dubokog učenja, a njihovi rezultati su evaluirani kako bi se pronašao model sa najvećom tačnošću.

Osnovni doprinosi ovog istraživanja mogu se sagledati iz tri različita aspekta. Prvi aspekt obuhvata uvođenje novog pristupa, zasnovanog na veštačkoj inteligenciji, za otkrivanje pokazatelja stresa usled suše kod biljaka suncokreta korišćenjem neinvazivnog snimanja, podataka iz više izvora i dubokog učenja. Drugi aspekt se odnosi na razvijanje prediktivnog modela za visinu prinosa suncokretovog ulja na osnovu različitih ekoloških i agronomskih faktora. Na kraju, istraživanje daje doprinos kroz kreiranje skupova podataka slika korenova i izdanaka suncokreta, kao i terenskih podataka o hibridima suncokreta u toku dve vegetacione sezone, što pruža podršku budućim istraživanjima u preciznoj poljoprivredi.

Ključne reči: veštačka inteligencija, duboko učenje, detekcija stresa usled suše, mašinsko učenje, predviđanje visine prinosa ulja, precizna poljoprivreda, slike iz rizotona, suncokret, transferno učenje.

Abstract

As the world population continues to grow, so does the demand for food and water, placing increasing pressure on the agricultural sector to enhance productivity while minimizing environmental impacts. Climate change plays a pivotal role in this process, as unpredictable weather patterns and prolonged droughts pose significant threats to crop yields and food security. In this context, there is a need to develop automated methods that enable agronomists to make informed and effective decisions, even under unfavorable climatic conditions, such as water scarcity combined with high temperatures. The development of artificial intelligence (AI)-based models offers a promising solution, enabling optimized resource management and improved resilience to variable weather conditions.

This research focuses on applying artificial intelligence technologies to optimize sunflower production by addressing two critical challenges: sunflower drought stress detection and sunflower oil yield prediction. Sunflower, as a major oilseed crop, is highly sensitive to water scarcity, which significantly impacts its growth and oil production. Traditional methods for assessing drought stress in sunflower plants involve a combination of physiological, technological, and molecular techniques. Similarly, conventional methods for estimating oil yield in new sunflower hybrids usually rely on field trials, chemical analyses, and statistical approaches. Although these methods provide reliable results, they are often time-consuming, costly, and require intensive human labor. To overcome these limitations, this study developed a methodology that integrates advanced machine and deep learning techniques with experimental and field data.

To detect drought stress, experiments were conducted on 12 sunflower inbred lines selected for their drought tolerance, productivity, and disease resistance. The plants were cultivated under controlled conditions in rhizotrons, enabling non-invasive monitoring of root system responses to varying soil moisture levels (70%, 50%, and 42% gravimetric water content). A high-resolution camera and specialized scanner were used to capture and analyze morphological changes in the root systems and shoots. Deep and transfer learning-based models were developed to detect drought stress using images of sunflower sections. Various model configurations were tested and fine-tuned to achieve the most accurate, reliable, and efficient model for stress detection.

To predict oil yield, field trials were conducted at four locations with diverse agro-ecological conditions in the Pannonian region of Serbia. Data from 1250

sunflower hybrids were used, with three replicates for each hybrid across all locations. The dataset included measurements of seed yield, oil content, and environmental factors such as temperature, precipitation, and soil properties, recorded over two growing seasons (2018 and 2019). Different machine and deep learning models were trained, and their performance was evaluated to identify the model with the highest accuracy.

The main contributions of this research include three key aspects. The first aspect involves the introduction of a novel AI-based approach for detecting drought stress in sunflower plants, leveraging non-invasive imaging, multisource datasets, and deep learning techniques. The second aspect focuses on the development of a predictive model for sunflower oil yield, incorporating various ecological and agronomic factors to ensure accurate and reliable forecasts. Finally, the research contributes through the creation of datasets of sunflower root and shoot images, as well as field data on sunflower hybrids collected over two growing seasons. These datasets serve as valuable resources to support future research in precision agriculture and crop optimization.

Key words: artificial intelligence, deep learning, drought stress detection, machine learning, oil yield prediction, precision agriculture, rhizotron imaging, sunflower, transfer learning.

1. Uvod

1.1 Motivacija

Jedan od glavnih izazova sa kojim se moderna poljoprivreda suočava su klimatske promene koje uključuju porast temperatura, promene u režimu padavina i učestalije pojave ekstremnih vremenskih uslova poput suša i poplava. Istovremeno, stalni rast broja stanovnika na globalnom nivou povećava potrebu za stabilnom, efikasnom i ubrzanom proizvodnjom hrane.

Pored toga, prinosi useva zavise od velikog broja drugih faktora, poput nivoa ugljen-dioksida, nivoa hranljivih materija u zemljištu, vlažnosti zemljišta i dostupnosti vode. Ovi faktori su međusobno povezani i promena bilo kog od njih utiče na strategiju obrade zemljišta. Takođe, usled klimatskih promena dolazi do porasta temperature vazduha i temperature vodenih površina, što može uzrokovati promenu staništa mnogih životinjskih vrsta, kao i druge poremećaje u ekosistemima.

Poljoprivreda koristi skoro 70% svih vodnih resursa širom sveta, što je čini najvećim potrošačem vode na globalnom nivou (Gleick, 2014). To znači da farmeri i agronomi ne samo da moraju da se prilagođavaju klimatskim promenama, već moraju i pažljivo da koriste vodne resurse u tom procesu. S obzirom na to da se poljoprivreda oslanja na prirodne resurse koji su sve više pod pritiskom (Arnell, 1999; Zhao et al., 2023; Farah et al., 2025), razvijanje održivih i efikasnih metoda proizvodnje postaje veoma značajno.

Klimatske promene, poput čestih sušnih perioda i ekstremnih vremenskih uslova, najviše pogađaju proizvodnju useva u regijama sa oskudnim vodnim resursima (Mbewu et al., 2024). Zbog toga je veoma važno pratiti potrošnju vode za svaki usev kako bi ona bila efikasno korišćena, osiguravajući održivo navodnjavanje i što veći prinos.

Suše su pojave koje se javljaju prirodno, ali klimatske promene su ih ubrzale i intenzivirale, sa mnogim posledicama kao što su povećan rizik po prirodu, živi svet i procese u poljoprivredi (Mukherjee et al., 2018). Posebno je zabrinjavajuće što se prema nekim istraživanjima i predviđanjima smatra da bi suša do 2050. godine mogla smanjiti prinose poljoprivrednih kultura za čak 50% (Li et al., 2009). Suša je jedan od najvažnijih abiotičkih faktora koji ograničava rast i razvoj uljanih biljnih vrsta, poput suncokreta, što može značajno da utiče na smanjenje njihovih prinosa.

Suncokret (*Helianthus annuus* L.) je veoma važna uljana biljna vrsta u industriji jestivog ulja i globalna potražnja za njim stalno raste. Agronomija se

konstantno bavi istraživanjem novih metoda koje bi dovele do povećanja proizvodnje suncokreta, uz istovremeno efikasno korišćenje zemljišnih i vodnih resursa (Cvejić et al., 2023). Jedan od ključnih pristupa povećanju proizvodnje je genetičko oplemenjivanje suncokreta.

Cilj savremenog oplemenjivanja suncokreta jeste stvaranje hibrida koji ostvaruju visok prinos čak i u uslovima izražene suše. Razvoj takvih hibrida zahteva duboko razumevanje genetičke varijabilnosti i adaptivnih mehanizama biljaka na vodni stres. Ključnu ulogu u ovom procesu ima fenotipska karakterizacija, pre svega kroz analizu morfoloških osobina biljaka, kao što su arhitektura biljke i dubina korena. U kombinaciji sa fiziološkim parametrima, poput efikasnosti korišćenja vode i fotosintetičke aktivnosti, fenotipska karakterizacija omogućava prepoznavanje genotipova sa većim potencijalom za toleranciju na sušu (Škorić et al., 2004; Nagel et al., 2012; Singh et al., 2018).

Tradicionalne metode za procenu stresa usled suše kod biljaka suncokreta obuhvataju kombinaciju fizioloških, biohemijskih i molekularnih metoda. Ove metode podrazumevaju pristupe kao što je analiza ekspresije gena koji učestvuju u odgovoru biljaka na sušni stres. U novije vreme, tradicionalne metode se dopunjuju savremenim tehnologijama kao što su senzori za vlagu u zemljištu, infracrvena termografija i multispektralne analize (Virnodkar et al., 2020; Kamarudin et al., 2021).

Tradicionalne metode za procenu visine prinosa ulja za nove hibride suncokreta zasnivaju se na poljskim ogledima, hemijskim analizama i statističkim metodama. To najčešće zahteva višegodišnja ispitivanja u različitim agroekološkim uslovima. Međutim, visoka varijabilnost spoljašnjih faktora, kao što su klimatske prilike (količina padavina i temperatura), može značajno uticati na prinos, što otežava procenu performansi novih hibrida (Andrianasolo et al., 2014; Călin et al., 2023).

Razvoj tehnologija poput mašinskog učenja (engl. *Machine Learning*, ML) i veštačke inteligencije (engl. *Artificial Intelligence*, AI), pruža automatizovanu analizu podataka, čime se unapređuju poljoprivredni procesi, uključujući metode optimizacije proizvodnje suncokreta (Crane-Droesch, 2018; Maffezzoli et al., 2022; Oliveira & Silva, 2023).

Rano identifikovanje reakcije biljaka na sušne periode može omogućiti prilagođavanje tehnika navodnjavanja i optimalno korišćenje dostupnih resursa, čime se može povećati otpornost useva na nepovoljne uslove. Sve ovo omogućava bolje donošenje odluka i održivo upravljanje resursima u uslovima suše. Korišćenjem ovih alata i velikih količina podataka može se

doprineti preciznijoj proceni prinosa i pravovremeno identifikovati reakcije biljaka na stres usled suše (Talaviya et al., 2020; Ali et al., 2024).

Ova disertacija će se baviti unapređenjem uzgoja suncokreta analizom slika i tabelarnih podataka, korišćenjem algoritama mašinskog i dubokog učenja (engl. *Deep Learning*, DL). Cilj predloženih metoda jeste razvoj sistema za detekciju stresa usled suše kod suncokreta (Lazić et al., 2024), kao i regresionih modela za procenu visine prinosa suncokretovog ulja (Cvejić et al., 2023). U odnosu na tradicionalne metode koje se oslanjaju na dugotrajne poljske ogledе, hemijske i laboratorijske analize, predloženi modeli omogućavaju bržu, automatizovanu i objektivniju procenu stanja useva.

1.2 Problem istraživanja

Problem istraživanja jeste nedostatak efikasnih i preciznih metoda za prepoznavanje stresa usled suše i predviđanje visine prinosa ulja kod suncokreta, što je dodatno otežano sve izraženijim efektima klimatskih promena.

Utvrđivanje stresa izazvanog sušom kod biljaka suncokreta se tradicionalno obavlja kombinovanjem fizioloških, molekularnih i biohemijskih metoda za praćenje reakcije biljke na sušu. Iako ove ustaljene metode pružaju korisne informacije, često nisu dovoljno precizne, zahtevaju puno vremena i nisu uvek lako primenljive u praksi.

Kada je reč o proceni visine prinosa ulja za nove hibride suncokreta, tradicionalno se koriste poljski ogledi, hemijske analize i statističke metode. Pored toga što ovaj proces zahteva višegodišnje eksperimente u različitim uslovima, on zahteva i morfološke i fiziološke analize koje zahtevaju veliko finansijsko ulaganje i mnogo vremena.

Još jedan problem tradicionalnih metoda je subjektivnost u vizuelnoj proceni biljaka, na koju se oslanjaju mnogi procesi, a koja otvara mogućnost za greške u proceni. Pored toga, tradicionalni statistički modeli imaju ograničenu sposobnost da obrade kompleksne međusobne odnose brojnih faktora kao što su vremenski uslovi i karakteristike zemljišta. Uključivanje savremenih tehnologija poput senzora za vlagu, infracrvene termografije i multispektralne analize u proces omogućava brže i pouzdanije praćenje stanja biljaka. Poteškoće u ovom procesu predstavljaju veliki troškovi i neophodnost organizovanja dodatnih tehničkih obuka za agronome koji se bave ovom oblašću.

Postoji potreba za razvojem automatizovanih metoda koje će agronomima omogućiti donošenje efikasnih odluka. Razvojem modela zasnovanih na veštačkoj inteligenciji mogli bi da se reše ovi problemi uz optimizovano korišćenje resursa u promenljivim klimatskim uslovima.

Savremene tehnologije bazirane na mašinskom i dubokom učenju suočavaju se sa značajnim izazovima među kojima je potreba za velikim, visokokvalitetnim skupovima podataka, koje je teško prikupiti u realnim uslovima. Pored toga, faktori kao što su tip tla, vremenske prilike i različitost uslova pod kojima su prikupljani podaci, mogu uticati na kvalitet podataka i otežati proces ekstrakcije obeležja. Modeli dubokog učenja se suočavaju i sa dodatnim problemima zato što zahtevaju značajne računarske resurse. Postoje i rizici preprilagođavanja (engl. *overfitting*) i nedostatka uvida u način na koji modeli donose odluke, jer se modeli dubokog učenja često posmatraju kao „crne kutije“.

U cilju prevazilaženja navedenih izazova sve više se koristi transferno učenje (engl. *Transfer Learning*, TL), koje koristi unapred obučene modele kojima treba manja količina podataka i koji smanjuju trajanje obuke. Dodatno, korišćenje objašnjive veštačke inteligencije (engl. *Explainable AI*) može doprineti razumevanju faktora koji utiču na odluke modela, što je značajno za transparentniju interpretaciju rezultata. Osim toga, korišćenje podataka iz različitih izvora, augmentacija, alteracija i sintetičko generisanje podataka doprinose poboljšanju performansi modela jer omogućavaju proširenje postojećeg skupa i smanjuju problem ograničenih resursa u pogledu sakupljenih podataka.

1.3 Predmet istraživanja

Predmet istraživanja je primena savremenih metoda veštačke inteligencije, posebno tehnika mašinskog i dubokog učenja, u analizi podataka o biljkama suncokreta radi detekcije stresa usled suše i predviđanja visine prinosa ulja.

Dakle, predmet istraživanja uključuje dva ključna aspekta:

1. detekciju stresa usled suše kod biljaka suncokreta na osnovu njihovih vizuelnih karakteristika,
2. predviđanje visine prinosa ulja suncokreta na osnovu podataka o hibridima i uslovima sredine.

1.3.1 Predmet istraživanja za detekciju stresa usled suše

Prvi zadatak ove doktorske disertacije podrazumeva primenu metoda dubokog učenja za automatsku detekciju stresa usled suše kod biljaka suncokreta na osnovu fotografija različitih delova suncokreta (koren, stablo, listovi).

Slike suncokreta korišćene u istraživanju su dobijene iz dva različita izvora:

1. fotografije korenovog sistema biljaka suncokreta iz rizotona (postavka koja omogućava posmatranje i proučavanje korenovog sistema biljaka u prirodnim ili kontrolisanim uslovima, bez vađenja biljke iz zemljišta),
2. digitalne slike delova suncokreta izvađenih iz zemlje (koren i nadzemni delovi biljke).

Važnu komponentu istraživanja predstavlja obogaćivanje podataka za kreiranje modele radi postizanja boljih rezultata. Ono će biti urađeno na dva različita načina:

1. korišćenjem standardnih metoda augmentacije, kao što su rotacija, translacija i promene kontrasta kod slika biljaka,
2. personalizovanim metodama relevantnim za specifične skupove podataka iz ovog istraživanja.

1.3.2 Predmet istraživanja za predviđanje visine prinosa ulja

Drugi zadatak ove doktorske disertacije podrazumeva primenu metoda mašinskog i dubokog učenja za predviđanje visine prinosa ulja suncokreta na osnovu numeričkih podataka o semenu suncokreta i lokacijama useva.

Ovaj deo istraživanja obuhvata implementaciju, evaluaciju i upoređivanje modela mašinskog učenja za predviđanje visine prinosa ulja, pri čemu se podaci analiziraju iz dva aspekta:

1. kroz agregirane podatke na nivou hibrida i lokacija, tako što će se za svaki hibrid posebno izračunati srednja vrednost prinosa na osnovu tri replikacije. Zatim će se za svaki hibrid izračunati prosečna vrednost prinosa na svim lokacijama,
2. kroz individualne performanse hibrida na svakoj lokaciji.

Ovakav pristup omogućava detaljniju analizu osobina hibrida i uticaja lokalnih uslova, što predstavlja metodološki doprinos istraživanju.

1.4 Ciljevi istraživanja

Opšti cilj ovog istraživanja je primena savremenih metoda veštačke inteligencije za unapređenje mogućnosti pravovremene detekcije stresa usled suše i predviđanja visine prinosa ulja kod suncokreta, čime se može doprineti održivoj poljoprivredi kroz efikasniju upotrebu resursa. Istraživanje ima dva specifična cilja:

1. razvoj modela za detekciju stresa usled suše kod biljaka suncokreta primenom mašinskog i dubokog učenja,
2. razvoj modela za predviđanje visine prinosa ulja suncokreta primenom mašinskog učenja.

1.4.1 Cilj istraživanja za detekciju stresa usled suše

Kada je reč o detekciji stresa usled suše kod biljaka suncokreta, cilj ovog istraživanja jeste razvoj modela dubokog učenja koji može da pruži uvid u tolerantnost inbred linija na sušu u kontrolisanim uslovima. Ovakav model bi u budućnosti mogao da se primeni kao alat u realnim poljoprivrednim uslovima za optimizaciju navodnjavanja i drugih faktora u zavisnosti od specifične otpornosti hibrida. Time bi se omogućilo agronomima da efikasnije koriste vodne resurse i smanje negativan uticaj na životnu sredinu.

Pored toga, cilj u istraživanju je ispitati u kojoj meri različite metode augmentacije slika mogu poboljšati generalizaciju modela za detekciju stresa izazvanog sušom.

1.4.2 Cilj istraživanja za predviđanje visine prinosa ulja

U slučaju predviđanja visine prinosa ulja, pored procene visine prinosa ulja, cilj je identifikovati najuticajnije faktore za predviđanje visine prinosa ulja kako bi se agronomima olakšao izbor optimalnih hibrida suncokreta.

Za razvoj modela mašinskog učenja korišćena su dva različita podskupa podataka. Korišćenjem prvog podskupa, fokus je stavljen na karakteristike hibrida, uključujući prinos semena, tip hibrida, otpornost na herbicide i parazite, i zrelost hibrida. U okviru ovog skupa podataka, koriste se srednje vrednosti sa više lokacija što omogućava određivanje ukupnih performansi različitih tipova hibrida. Drugim podskupom obuhvaćeni su podaci sa svake lokacije (poput osobina lokacija i vremenskih uslova) i na taj način se analizira uticaj ekoloških faktora na visinu prinosa ulja.

1.5 Hipoteze u istraživanju

1.5.1 Hipoteze za detekciju stresa usled suše

H1: Primenom dubokog učenja na slike korena i nadzemnih delova suncokreta gajenog u rizotronima, može se identifikovati stres usled suše.

H2: Postoji značajna korelacija između morfoloških karakteristika korena i nadzemnih delova suncokreta gajenog u rizotronima i nivoa stresa usled suše.

H3: Obogaćivanje slika (augmentacijom i alternacijom) može poboljšati performanse DL modela u detekciji stresa usled suše na ograničenim skupovima podataka kod suncokreta gajenog u rizotronima.

1.5.2 Hipoteze za predviđanje visine prinosa ulja

H4: Modeli bazirani na mašinskom učenju omogućavaju procenu prinosa suncokretovog ulja na osnovu genetičkih karakteristika hibrida (tip ulja, otpornost na herbicide, sadržaj oleinske kiseline, tip konditorskih proizvoda, rezistencija na volovod i plamenjaču, zrelost hibrida i sadržaj ulja) i lokacija useva (tip zemljišta i meteorološki podaci).

H5: Postoji značajna korelacija između genetičkih karakteristika hibrida (tip ulja, otpornost na herbicide, sadržaj oleinske kiseline, tip konditorskih proizvoda, rezistencija na volovod i plamenjaču, zrelost hibrida i sadržaj ulja) suncokreta i prinosa ulja, što omogućava selekciju optimalnih hibrida za različite klimatske uslove.

1.6 Struktura teze

Ova disertacija je organizovana u sedam poglavlja.

Prvo poglavlje daje kratak osvrt na kontekst iz kog je proisteklo ovo istraživanje. Objašnjeni su motivacija, problem i predmet istraživanja kroz dva različita zadatka. Takođe su uvedeni ciljevi i hipoteze istraživanja.

U drugom poglavlju nalazi se pregled aktuelnog stanja iz oblasti ove doktorske disertacije. Pregled, pored opštog konteksta primene veštačke inteligencije u agronomiji, obuhvata i detaljnu analizu postojećih rešenja za svaki od zadataka. Opisani su i specifični pristupi transfernog učenja za domen specifične podatke, kao i sistematizacija ključnih problema i njihovih

potencijalnih rešenja sa posebnim osvrtom na aktuelna istraživanja relevantna za oba zadatka koji su predmet ove disertacije.

Treće poglavlje predstavlja teorijske osnove istraživanja, počevši od opštih koncepata veštačke inteligencije, mašinskog i dubokog učenja koji čine širi okvir za razumevanje ove doktorske disertacije. Posebna pažnja posvećena je algoritmima mašinskog učenja, konvolucionim neuronskim mrežama i transfernom učenju kao ključnoj tehnici za analizu slika kada se radi sa ograničenim skupovima podataka.

Četvrto poglavlje se bavi detaljnim opisom metodologije i eksperimentalne postavke za oba zadatka iz ovog istraživanja. U okviru poglavlja opisani su eksperimenti, skupovi podataka i njihova obrada i analiza. Takođe su obrazloženi razlozi odabira korišćenih modela, nakon čega je istaknut način pripremanja podataka za modele. Na kraju su opisane metode evaluacije koje su korišćene za procenu uspešnosti modela.

U petom poglavlju je data sveobuhvatna evaluacija performansi razvijenih modela za zadatke disertacije. Detaljno su analizirani rezultati dobijeni primenom odabranih metrika, sa posebnim osvrtom na modele koji su tokom preliminarnih ispitivanja pokazali najbolje performanse.

Šesto poglavlje sadrži diskusiju rezultata i analizu grešaka kreiranih modela. Poglavlje opisuje kako se dobijeni rezultati odnose na hipoteze iz uvoda. Pored toga, poglavlje sadrži poređenje dobijenih rezultata sa rezultatima modela koji se bave rešavanjem sličnih zadataka. Takođe, napisana je i diskusija o ograničenjima ovog istraživanja.

Na kraju su u sedmom poglavlju diskutovani izvedeni zaključci i predložene smernice za nastavak istraživanja i budući rad.

2. Pregled literature i aktuelno stanje u oblasti

U ovom poglavlju dat je pregled aktuelnog stanja u oblasti primene veštačke inteligencije u agronomiji, sa posebnim osvrtom na detekciju stresa usled suše i predviđanje visine prinosa. Prikazani su savremeni pristupi, korišćeni podaci i algoritmi, kao i njihova ograničenja. Na kraju poglavlja analizirani su ključni problemi u primeni ovih metoda i razmatrana potencijalna rešenja.

2.1 Primena veštačke inteligencije u agronomiji

Tehnologije zasnovane na veštačkoj inteligenciji doprinose poboljšanju efikasnosti u brojnim oblastima poljoprivrede, uključujući oblasti koje se bave navodnjavanjem, praćenjem sadržaja vode u zemljištu, postavljanjem useva i praćenjem njihovog prinosa, suzbijanjem korova i detekcijom bolesti biljaka (Talaviya et al., 2020). U okviru navedene široke primene veštačke inteligencije u poljoprivredi, ističu se brojne studije koje se bave prepoznavanjem bolesti biljaka na osnovu njihovih slika (Yuan et al., 2022), IoT (engl. *Internet of Things*) rešenjima za pametnu poljoprivredu (Farooq et al., 2020), velikim podacima (engl. *big data*) u poljoprivredi (Wolfert et al., 2017), kao i konceptom poljoprivrede 4.0 (Maffezzoli et al., 2022; Araújo et al., 2021; de Oliveira & de S. e Silva, 2023).

Za primenu veštačke inteligencije u agronomiji često se koriste podaci koji su dobijeni daljinskom detekcijom u kombinaciji sa senzorskim tehnologijama. Daljinska detekcija igra značajnu ulogu u praćenju velikih poljoprivrednih površina, što se može postići korišćenjem satelitskih vegetacionih indeksa kao što je NDVI (engl. *Normalized Difference Vegetation Index*) (Pettorelli et al., 2005; Mulla, 2013; Xue & Su, 2017). Pored toga, senzori mogu da se koriste za prikupljanje podataka u realnom vremenu kao što su, na primer, podaci o vlažnosti tla, temperaturi i vlažnosti vazduha (Mittelbach et al., 2011; Gutiérrez et al., 2014). Ovi podaci se u kombinaciji sa istorijskim meteorološkim podacima mogu koristiti za kreiranje AI modela koji donose optimalne odluke o navodnjavanju. Na taj način, ovakvi sistemi omogućavaju efikasnije upravljanje resursima i prilagođavanje promenljivim uslovima okruženja (Hussain et al., 2024).

Sistemi za pomoć donošenju odluka (engl. *Decision Support Systems*) (Rose et al., 2016; Lindblom et al., 2017) koriste se za analizu podataka, izvođenje zaključaka i upravljanje procesima, s ciljem ubrzanja i unapređivanja celokupnog procesa. Ovi sistemi povećavaju efikasnost i nalaze široku primenu u različitim aspektima poljoprivredne proizvodnje. Jedan od značajnih primera njihove primene je u optimalnom upravljanju vodom za navodnjavanje, gde se koriste podaci o potrebama za navodnjavanje, stvarnim

zalihama vode i drugim lokalnim faktorima (Navarro-Hellín et al., 2016; Viani et al., 2017; Karim et al., 2017).

Često se koriste i podaci dobijeni daljinskom detekcijom iz više izvora, što može da omogući izgradnju sveobuhvatnih modela za praćenje različitih procesa u poljoprivredi. U tom kontekstu, Shen et al. (2019) primenjuju kombinaciju podataka sa dva različita satelitska sistema za praćenje suše u provinciji Henan u Kini. Jedan sistem pruža podatke o vegetaciji, temperaturi površine zemljišta i indeksima suše, a drugi meri količinu padavina i prati hidrološke i meteorološke uslove. Na osnovu rezultata istraživanja u Kini, vidi se da je moguće obezbediti efikasno praćenje meteorološke i poljoprivredne suše. Ujedno, zaključeno je da je postignuta snažna korelacija između suše i meteoroloških indeksa na lokalnom nivou.

Metoda koja je slična daljinskoj detekciji, odnosno termovizija zasnovana na dronovima (engl. *Unmanned Aerial Vehicle*, UAV) pokazala se takođe korisnom za prikupljanje poljoprivrednih podataka. Korišćenjem termovizije precizno se prikupljaju termalni podaci o usevima, a prikupljanjem ovih podataka u realnom vremenu mogu se otkriti temperaturne varijacije koje su povezane sa vodnim stresom. Analizom termalnih podataka u realnom vremenu izvršena je procena biomase i prinosa 18 genotipova pšenice izloženih različitim nivoima vodnog stresa u agroekološkim uslovima Australije. Kombinovanjem termalnih podataka sa naprednim metodama mašinskog učenja identifikovani su ključni indeksi vodnog stresa, kao i genotipovi sa povećanom produktivnošću na različitim tipovima zemljišta (Das et al., 2021).

Od kada su algoritmi dubokog učenja počeli intenzivno da se razvijaju, beleženi su izuzetni rezultati do kojih su doveli u brojnim oblastima. Široka primena i efikasnost ovih metoda podsticali su naučnu zajednicu da ih testira i u novim kontekstima. Poslednjih godina postoji sve više istraživanja koja se bave primenom dubokog učenja u preciznoj poljoprivredi. Specifično, upotrebom konvolucionih neuronskih mreža došlo je do značajnih promena u zadacima obrade i analize slika u poljoprivredi.

Danas postoji široka upotreba modela dubokog učenja u zadacima kao što su segmentacija biljaka i njihovih delova, detekcija i prepoznavanje bolesti biljaka, klasifikacija biljaka po različitim kriterijumima i slično. Kada je reč o klasifikaciji stresa usled suše, modeli dubokog učenja su upotrebljavani na različitim biljnim kulturama, među kojima je i kukuruz (An et al., 2019). U navedenom istraživanju stres usled suše kod kukuruza klasifikovan je primenom duboke konvolucione neuronske mreže na RGB slikama listova,

pri čemu su biljke, na osnovu režima navodnjavanja i vizuelno ispoljenih simptoma, svrstane u više kategorija intenziteta sušnog stresa.

Transferno učenje je značajno doprinelo uspešnoj primeni dubokih neuronskih mreža (Pan & Yang, 2010; Yosinski et al., 2014). Ono omogućava prilagođavanje unapred obučених neuronskih mreža specifičnim zadacima. Kroz fino podešavanje mreža na specifičnim skupovima podataka, moguće je postići visoku tačnost predikcija čak i u slučaju malih skupova podataka (Shin et al., 2016), koji su česta pojava u realnom svetu. Uspešna primena TL u otkrivanju vodnog stresa kod različitih useva, uključujući kukuruz, pokazala je da su ovi modeli precizniji, neinvazivniji, ekonomičniji i skalabilniji u poređenju sa tradicionalnim metodama detekcije (Virnodkar et al., 2020; Al Sahili & Awad, 2022).

2.2 Primena veštačke inteligencije u detekciji stresa usled suše

Tokom poslednjih godina se primećuje porast broja istraživanja u vezi sa primenom veštačke inteligencije na detekciju stresa usled suše kod mnogih biljaka. Ove tehnologije primenjuju se za pravovremenu detekciju stresa usled suše, što je ključno za optimizaciju navodnjavanja, očuvanje resursa i povećanje produktivnosti u poljoprivredi.

Tradicionalne metode praćenja stresa usled suše obuhvataju neposrednu analizu klasičnih meteoroloških podataka i procenu prinosa useva posmatrajući predašnje rezultate i iskustvo. Problem kod ovih metoda je to što su vremenski i finansijski zahtevne. Rešenja bazirana na veštačkoj inteligenciji pružaju potporu bržem, preciznijem i skalabilnijem rešenju. Kombinacija klasičnih laboratorijskih analiza i tehnika veštačke inteligencije može pružiti širi uvid u kompleksne reakcije biljaka na stresne uslove, uključujući stres usled suše. Ovo može doprineti boljem razumevanju interakcije biljke i okruženja, što je posebno značajno za unapređivanje postojećih modela za detekciju stresa usled suše (Ali et al., 2024).

Jedan značajan iskorak u odnosu na tradicionalne prakse ostvaren je spajanjem tehnika mašinskog učenja i računarske obrade slike u cilju precizne procene vodnog stresa kod biljaka. Na ovaj način može se unaprediti praćenje reakcije biljaka na sušu na automatski, brz i neinvazivan način (Chandel et al., 2021). Za ovo se najčešće koriste podaci dobijeni daljinskom detekcijom uz obradu slika konvolucionim neuronskim mrežama (engl. *Convolutional Neural Network*, CNN).

ML metode su se pokazale kao pouzdane za klasifikaciju stresa usled suše i uz pomoć spektra refleksije listova. Rezultati studije koja se bavila time

pokazuju potencijal spektralne analize u kombinaciji sa naprednim računarskim tehnikama za pravovremeno prepoznavanje reakcija biljaka na sušu (Barradas et al., 2021).

Podaci dobijeni daljinskom detekcijom, uz pomoć senzora i bespilotnih letelica, često se koriste u kombinaciji sa metodama veštačke inteligencije radi praćenja stresa usled suše (Virnodkar et al., 2020; Suharso et al., 2022; Ali et al., 2024; Sharma et al., 2025). Podaci koji se najčešće koriste obuhvataju informacije o specifičnim faktorima, kao što su tip zemljišta, klima i karakteristike useva. Na osnovu navedenog se vidi da korišćenje veštačke inteligencije za procenu, praćenje i predviđanje suše pokazuje brz rast i tendenciju da se njen uticaj u budućnosti poveća (Kikon & Deka, 2022).

Mašinsko učenje i duboko učenje su se pokazali kao moćni alati za predviđanje stresa usled suše kod raznih poljoprivrednih kultura (tabela 2), nudeći inovativna rešenja analizom fenotipskih promena u usevima, kao što su boja, tekstura i oblik lista. Karakteristike slika, kao što su boja i tekstura, najviše se izdvajaju i koriste za identifikaciju i klasifikaciju nivoa stresa usled suše. Na ovaj način je omogućena automatizovana, nedestruktivna i efikasna analiza, čime se smanjuje potreba za ručnim posmatranjem.

Pored toga, transferno učenje omogućava modelima koji su prethodno obučeni na velikim skupovima podataka da se prilagode novim, manjim skupovima podataka kakvi se često susreću u praksi. Na taj način se ubrzava konvergencija modela i poboljšava njegova preciznost i generalizacija. Na primer, modeli obučeni na slikama kao što je *ImageNet* (Deng et al., 2009) mogu se fino podesiti za zadatke klasifikacije medicinskih slika sa mnogo manjim skupovima podataka (Shin et al., 2016).

Prilagođeni CNN okviri za detekciju stresa usled suše u usevima kukuruza su u kombinaciji sa transfernim učenjem pokazali superiorne performanse, sa specifičnim arhitekturama konvolucionih neuronskih mreža, kao što su ResNet50 i EfficientNetB1, koje su postigle preko 99% tačnosti (Goyal et al., 2024). Ovaj primer ističe potencijal ovog pristupa za rešavanje kritičnih poljoprivrednih izazova korišćenjem strategija za povećanje trening skupa, prilagođavanja domena i unapred obučениh modela.

U cilju boljeg razumevanja postojećih pristupa i metoda za detekciju biljnog stresa, u nastavku su u obliku tabela prikazani relevantni naučni radovi koji se bave identifikacijom dva ključna tipa stresa. Najpre je, u tabeli 1, prikazan pregled istraživanja iz oblasti biotičkog stresa (stresa izazvanog patogenima). U tabeli se nalaze informacije o biljnim vrstama korišćenim u istraživanju, vrsti algoritama koji su korišćeni za predikciju, kao i najbolji rezultat koji je postignut u istraživanju. Tabela 2 predstavlja pregled najznačajnijih

istraživanja u oblastima predikcije sušnog stresa (abiotički stres), sa posebnim osvrtom na primenjene algoritme, ostvarene rezultate i ograničenja korišćenog pristupa.

Tabela 1. Pregled relevantnih radova za detekciju biotičkog stresa.

Istraživanje	Biljna vrsta	Tip algoritma	Performanse
Amara et al., 2017	Banana	LeNet	92,88% (tačnost, slike u boji) 85,94% (tačnost, sivi tonovi)
Liu et al., 2017	Jabuka	AlexNet, GoogLeNet, VGGNet-16, ResNet-20	97,62% (tačnost)
Ramcharan et al., 2017	Manioka	Inception-v3, ImageNet	93% (tačnost)
Ferentinos, 2018		AlexNet, AlexNetOWTBn, GoogLeNet, Overfeat, VGG	99,53% (tačnost)
Mohanty et al., 2016	Biljke iz PlantVillage skupa podataka	AlexNet, GoogLeNet	99% (tačnost, original) 98% (tačnost, sive nijanse)
Cruz et al., 2017	Maslina	Modified LeNet	98,6% (tačnost)
Fujita et al., 2016	Krastavac	CNN	82,3% (tačnost)
Sladojevic et al., 2016	Kruška, višnja, breskva, jabuka, grožđe	CaffeNet, ImageNet	96,3% (tačnost)
Fuentes et al., 2017	Paradajz	AlexNet, ZFNet, VGG-16, GoogLeNet, ResNet-50, ResNet-101, ResNetXt-101, Faster RCNN, R-FCN, SSD	55,64% (tačnost, bez augmentacije podataka) 83,06% (tačnost, sa augmentacijom podataka)
Lu et al., 2017	Pšenica	VGG-FCN, VGG-CNN	97,95% (tačnost)

Tabela 2. Pregled relevantnih radova za detekciju stresa usled suše.

Istraživanje	Tip algoritma	Performanse	Ograničenja
Barradas et al., 2021	RandomForest	94% (tačnost)	Varijabilnost za druge biljke i uslove životne sredine upitna
Goyal et al., 2024	ResNet50, EfficientNetB1	99,53% (tačnost)	Mali i jednorodni skup podataka prikupljen iz jednog eksperimenta, što otežava varijabilnost
Kuo et al., 2023	ResGCNet	90% (tačnost)	Tačnost značajno

			opada sa manjim brojem ulaznih atributa
Hagen et al., 2024	Random forest, RNA	92,3% (tačnost)	Varijabilnost za druge biljke i uslove životne sredine upitna
Chandel et al., 2021	CNN, DCNN, GoogLeNet	98,3% (tačnost)	Performanse se možda ne mogu generalizovati
Swaminathan & Vairavasundaram, 2024	CNN	95,42% (tačnost)	Uslovi životne sredine nisu dovoljno uključeni u ulaz modela
Azimi et al., 2021	BAT-optimized ResNet-18	96% (tačnost)	Nije korišćen javno dostupan skup podataka, već skup podataka usko vezan za opisane uslove
Rojanarungruengporn & Pumrin, 2021	VGG16+LSTM	85% (tačnost)	Nisu pokriveni izazovi u vezi sa primenom predloženog modela u realnim uslovima
Chandel et al., 2022	CNN, AlexNet, GoogLeNet, Inception v1, Inception v3, MobileNet v2, ResNet50	98,4% (tačnost)	Varijabilnost za druge biljke i uslove životne sredine upitna
Concepcion Li et al., 2020	MobileNetV2, ResNet101, InceptionV3	0,6568 (R^2)	Varijabilnost za druge biljke i uslove životne sredine upitna

Korišćenje podataka koji su prikupljeni uz pomoć senzora i daljinskog istraživanja zajedno sa tehnologijama mašinskog učenja omogućava preciznu, brzu i široko primenjivu detekciju različitih vrsta stresa kod useva. Činjenica da ovakva metodologija omogućava primenu rezultata u realnom vremenu olakšava primenu ciljanih strategija upravljanja.

Međutim, važno je napomenuti da uključivanje podataka dobijenih daljinskim putem često zahteva značajna finansijska ulaganja. Stoga je pre ulaganja u satelitske snimke ili srodne tehnologije, značajno sprovesti preliminarne eksperimente kako bi se procenilo da li se kompleksniji pristupi automatske daljinske detekcije mogu efikasno primeniti na konkretan problem. U naše istraživanje neće biti uključeni podaci daljinskog istraživanja, već isključivo podaci prikupljeni u kontrolisanim uslovima kroz eksperimente u rizotronima – specijalizovanim providnim laboratorijama od stakla koje omogućavaju nedestruktivna, ponovljena posmatranja i merenja korenovog sistema.

2.3 Primena veštačke inteligencije u predviđanju visine prinosa ulja

Jedan od najznačajnijih ciljeva u preciznoj poljoprivredi je poboljšanje proizvodnje i kvaliteta prinosa uz minimizovanje troškova (Zhang et al., 2002; Gebbers & Adamchuk, 2010). Uzimajući ovo u obzir, očekivano je da postoji veliki broj istraživanja koji se bavi pravovremenim procenama prinosa različitih poljoprivrednih kultura. Drugi zadatak ove disertacije je predikcija prinosa suncokreta čije je ulje važan prehrambeni resurs na globalnom nivou.

Rana procena prinosa suncokretovog ulja značajna je zato što može pomoći u procesu selekcije suncokreta (genotipskih varijanti suncokreta sa visokim prinosom ulja). Prinos suncokreta zavisi od više različitih karakteristika među kojima su genetičke i fenotipske karakteristike hibrida, vremenski uslovi i fizičko-hemijske karakteristike zemljišta. Pravovremena i precizna selekcija suncokreta može da dovede do skraćivanja oplemenjivačkih planova.

Poslednjih godina broj istraživanja u kojima se ML algoritmi koriste za predviđanje prinosa semena i ulja suncokreta je u porastu. Primena ovih modela je pretežno usmerena na predviđanje ishoda prinosa (Khalifani et al., 2022; Călin et al., 2023), kao i na predviđanje vizuelnih karakteristika koje pomažu u identifikaciji bolesti i rezistencije na bolesti koje su često značajno povezane sa prinosom (Malik et al., 2022; Ghosh et al., 2023). Takođe, značajan broj studija bavi se procenom kvaliteta semena (Çetin et al., 2021; Lužaić et al., 2023) čime je potpomognuta selekcija visoko produktivnih hibrida suncokreta. Badshah et al. (2024) predložili su model klasifikacije koji koristi slučajnu šumu koja sa tačnošću od 99,7% omogućava preporuku najprikladnijih useva na osnovu parametara zemljišta i uslova životne sredine. Ovi rezultati ukazuju na potencijal algoritama mašinskog učenja da podrži automatizaciju upravljanja proizvodnjom suncokreta i olakša selekciju najboljih hibrida u skladu sa specifičnim agroekološkim uslovima.

Potreba za pravovremenim otkrivanjem međusobnih relacija karakteristika od kojih zavisi prinos ulja, dovela je do usvajanja tehnologija daljinske detekcije i senzora (Campbell et al., 2024) u preciznoj poljoprivredi (Curran, 1987). Modeli mašinskog učenja se koriste za analizu slika suncokreta dobijenih daljinskim putem. Većina ovakvih istraživanja oslanja se na metapodatke za kreiranje modela predikcije prinosa ili bolesti biljaka, uzimajući u obzir različite parametre poput vremenskih prilika i osobina zemljišta (Kurtulmuş, 2021; Amankulova et al., 2023; Gulzar et al., 2023). Ovi modeli su korišćeni za prognozu prinosa ili identifikaciju i dijagnostikovanje bolesti suncokreta, na osnovu vizuelnih simptoma ili numeričkih podataka o biljkama dobijenih u realnom vremenu putem senzora, što omogućava pravovremeno i efikasno upravljanje radi očuvanja useva.

Pored toga, tehnike mašinskog učenja korišćene su za precizno predviđanje prinosa pšenice, kukuruza i pamuka koristeći vegetacione indekse izvedene iz *Landsat* satelitskih podataka. Satir & Berberoglu (2016) koriste *Stepwise Linear Regression* algoritam za procenu prinosa koji je postigao visoku tačnost. Za pšenicu je prosečna greška bila oko 7,9%, za kukuruz oko 8,8%, dok je za pamuk greška bila najmanja – oko 6,3%.

Duboko učenje u kombinaciji sa pojačanim učenjem (engl. *Reinforcement Learning*) može rešiti ograničenja postojećih modela predviđanja prinosa tako što omogućava direktno mapiranje između neobrađenih podataka i vrednosti prinosa. Jedan od primera za to je istraživanje koje koristi model Deep Recurrent Q-Network, koji kombinuje rekurentu neuronsku mrežu i Q-Learning algoritam za kreiranje okvira za predviđanje prinosa useva. Ovaj model mapirajući neobrađene podatke na vrednosti prinosa postiže tačnost od 93.7%, čime prevazilazi ograničenja postojećih metoda i obezbeđuje očuvanje distribucije originalnih podataka (Elavarasan & Vincent, 2020).

Još jedno relevantno istraživanje je primena statističkih modela za predviđanje koncentracije ulja u semenu suncokreta na osnovu podataka iz eksperimenata sprovedenih na 61 sorti u Francuskoj između 2000. i 2011. godine (Andrianasolo et al., 2014). Korišćeno je 25 prediktora za izgradnju tri modela (višestruka linearna regresija, generalizovani aditivni model – GAM i regresiono stablo), pri čemu je GAM model postigao najbolju preciznost (RMSEP = 1,95%; EF = 0,71). Ključni prediktori su bili: potencijalna koncentracija ulja u zrnu (OC), pokazatelji funkcionalnosti lisne mase nakon cvetanja (LAD2 i MRUE2), kao i parametri koji ukazuju na ishranu biljaka azotom i njihov vodni status. Ovi parametri su omogućili bolje razumevanje uticaja ishrane i vodosnabdevanja na formiranje ulja u semenu, čime je unapređen model za tačno predviđanje koncentracije ulja kod različitih genotipova.

U cilju boljeg razumevanja postojećih pristupa i metoda za predikciju prinosa različitih biljnih kultura, u nastavku su u obliku tabele prikazani relevantni naučni radovi koji se time bave. U tabeli 3 se nalaze informacije o biljnim vrstama korišćenim u istraživanju, obeležjima za predikciju, vrsti algoritma koji je korišćen za predikciju i najboljem rezultatu koji je postignut u istraživanju.

Tabela 3. Pregled relevantnih radova za predviđanje prinosa.

Istraživanje	Biljna vrsta	Karakteristike za predikciju	Algoritam	Performanse
Shahhosseini et al., 2020	Kukuruz	Podaci o upravljanju usevima, podaci o klimi, osobine	LR, LASSO, XGBoost, lightGBM,	9,56% (RRMSE optimizovanog

		zemljišta, istorijski podaci o prinosu	RF	ponderisanog ansambla)
Panda et al., 2010	Kukuruz	Informacije o navodnjavanju, osobine zemljišta	BPNN	96,04% (RMSE)
Crane-Droesch, 2018	Kukuruz	Klimatske informacije, osobine zemljišta, informacije o navodnjavanju	SNN, OLS regresija	251,5 (MSE)
Khaki and Wang, 2019	Kukuruz	Genotip useva, prinosi, vremenski i zemljišni uslovi	DNN, LASSO, SNN, RT	RMSE: 12% prosečnog prinosa (50% SD)
Zang et al., 2020	Kukuruz	Informacije o navodnjavanju, osobine zemljišta, klimatske informacije, satelitske varijable, informacije o obradivim površinama, SIF podaci	LASSO, RF, XGBoost, LSTM	0,77 (R^2 (SIF))
Khaki et al., 2020	Kukuruz i soja	Istorijski podaci o prinosu, vremenski parametri, karakteristike zemljišta, podaci o upravljanju usevima	CNN-RNN model	Postignuto poboljšanje predikcije prinosa u odnosu na osnovne modele
Peng et al., 2020	Kukuruz i soja	Istorijski podaci o prinosu, klimatske informacije, MODIS podaci, satelitski podaci o fluorescenciji hlorofila (SIF)	LASSO, RIDGE, SVR, RF, ANN	Kukuruz: 18,11, Soja: 5,31 (RMSE)
Maya Gopal & Bhargavi, 2019	Neoljušteni pirinač	Navodnjavanje, obradive površine, podaci o prinosu, upravljanje usevima, đubrenje	ANN, SVR, K-NN, RF	RF: 0,085 (RMSE)
Maya Gopal & Bhargavi, 2019	Neoljušteni pirinač	Informacije o obradivim površinama, klimatske informacije, osobine zemljišta, podaci o poljoprivrednoj proizvodnji, informacije o navodnjavanju	MLR-ANN model	RMSE: 0,051, MAE: 0,041, R^2 : 0,99
Elavarasan et al., 2020	Neoljušteni pirinač	Klimatske informacije, produktivnost zemljišta, karakteristike i dostupnost podzemnih voda	RF, DT, Gradient Boosting	91,23% (tačnost)
Guo et al., 2021	Pirinač	Fenologija, klimatske informacije i informacije o obradivim	BPNN, RF, SVM	SVM: 0,33 (R^2)

		površinama		
Sun et al., 2019	Soja	Istorijski podaci o prinosu, klimatske informacije, MODIS podaci, informacije o obradivim površinama	CNN-LSTM	0,78 (R ²)
Yalcin, 2019	Suncokret	Slike sa terena, informacije o rastu useva	CNN	87,67% (tačnost za validacioni skup)
Lin et al., 2019	Paradajz	Klimatske informacije, informacije o rastu useva	TOMGRO i Vanthoor model	2,5974 (RMSE)
Abbas et al., 2020	Krompir	Provodljivost zemljišta, vlažnost, nagib, NDVI, hemijski sastav	LR, Elastic Net, K-NN, SVR	SVR: 4,62 (RMSE)
Guo & Xue, 2014	Pšenica	Informacije o obradivim površinama	ANN (MLP)	0,0001 (MSE)
Zang et al., 2020	Pšenica	Klima, zemljište, vegetacioni indeks, obradive površine, podaci o prinosu	SVM, GPR, RF, DT, NN, K-NN, BST, BGT	RF: 0,81 (R ²)
Cao et al., 2020	Pšenica	Informacije o usevu, podaci o prinosu, osobine zemljišta, socioekonomski faktori, klimatske informacije, vegetacioni indeksi	Ridge, RF, blagi GBM	0,68–0,75 (R ²)
Sánchez et al., 2014	Razne biljke	Klima, informacije o obradivim površinama, informacije o navodnjavanju	MLR, DT, NN, SVR, K-NN	SVR: 4,91 (RMSE)

Pregled postojećih istraživanja pokazuje da se za predikciju prinosa najčešće koriste kombinacije klimatskih informacija, osobina zemljišta, podataka o navodnjavanju, istorijskih prinosa, kao i informacija o upravljanju usevima. Među najzastupljenijim vrstama u istraživanjima dominiraju kukuruz, pirinač i pšenica, dok su soja, krompir, paradajz i druge biljke analizirane u manjem broju radova. Većina radova koristi numeričke i tabularne podatke za ulaz u model.

Inovacija ovog rada u odnosu na ranija istraživanja predikcije prinosa je u tome što se koriste podaci koji obuhvataju srednje vrednosti prinosa po lokacijama i hibridima, kao i individualne performanse hibrida na svakoj lokaciji. Implementirani modeli se zasnivaju na analizi genetički indukovanih karakteristika hibrida (kao što su tip ulja, sadržaj oleinske kiseline, otpornost na herbicide i parazite, i zrelost) i specifičnostima lokacija (meteorološki

podaci i tip zemljišta), što omogućava preciznu procenu prinosa i selekciju optimalnih hibrida za različite klimatske uslove.

2.4 Problemi i njihova potencijalna rešenja u primeni veštačke inteligencije u agronomiji

Iako uključivanje savremenih metoda veštačke inteligencije može imati pozitivan uticaj na agronomiju, postoje različiti izazovi koji mogu ugroziti efikasnost rezultata.

Na primer, za uspešnu primenu dubokog učenja u agronomiji potrebni su veliki i kvalitetni skupovi podataka kako bi modeli postigli visoku tačnost (Halevy et al., 2009; Leminen Madsen et al., 2020). Ovaj problem nije prisutan samo u slučaju agronomskih podataka, već je i generalan problem kod dubokog učenja. Važno je da se obezbede visokokvalitetni, standardizovani podaci iz različitih izvora, koji često koriste različite formate i protokole, zbog čega mogu da budu teški za integraciju u AI modele. Pored toga, potrebno je obezbediti dovoljno podataka za validaciju modela, ali i osigurati odgovorno korišćenje podataka (Williamson et al., 2023).

Jedan od problema dubokog učenja je problem prilagođavanja. Ovaj problem se javlja tokom faze obuke i dovodi do visoke varijanse modela. Veštačkim povećanjem raznolikosti trening slika može se poboljšati generalizacija modela. Povećanje raznolikosti trening slika može se postići augmentacijom i sintetičkim generisanjem novih primeraka na osnovu postojećih podataka (Douarre et al., 2018).

Pošto je količina dostupnih podataka u scenarijima iz realnog sveta ograničena, uobičajeno je da se za klasifikaciju slika koriste lagane CNN arhitekture u kombinaciji sa transfernim učenjem. Korišćenje transfernog učenja može poboljšati tačnost modela bez potrebe za dodatnim podacima (Al Sahili & Awad, 2022). Ovakvi modeli su pokazali obećavajući balans između visoke tačnosti klasifikacije i računarske efikasnosti u mnogim oblastima (Zhang et al., 2020).

Na primer, pokazalo se da transferno učenje olakšava dijagnozu bolesti biljaka, omogućavajući modelima obučanim na velikim skupovima podataka da se prilagode manjim skupovima podataka specifičnim za domen. Primer jednog takvog istraživanja je identifikacija bolesti plesni u bisernom prosu primenom transfernog učenja, gde su postignuti izuzetni rezultati, odnosno tačnost od 95% (Coulibaly et al., 2019). Izazov kod korišćenja transfernog učenja jeste što su slike poljoprivrednih kultura domen specifične, a

prethodno obučeni modeli koji se koriste u TL-u nisu obučeni na takvim slikama, već na opštim skupovima podataka slika.

Još jedan izazov koji se često javlja kod agronomskih podataka može biti neravnomerna distribucija klasa (broj instanci u grupama), što takođe može degradirati performanse DL modela. Na primer, to se vidi u radu koji predstavlja primenu TL kroz ResNet50 arhitekturu za detekciju uobičajenih bolesti jabuke. Dobijena ukupna tačnost je bila 97%, ali je tačnost za jednu od kategorija bila samo 51%, što je uzrokovano neravnotežom klasa (Thapa et al., 2020). Da bi se prevazišao ovaj problem, često se koriste tehnike obogaćivanja podataka i balansiranja klasa. Povećanje količine podataka je idealna opcija, ali je to često vremenski i finansijski zahtevno ili čak neizvodljivo u datom trenutku. Alternativno rešenje može biti korišćenje sintetički generisanih podataka, koji omogućavaju potencijalno prevazilaženje tih ograničenja (Yang et al., 2022).

Količina korisnih informacija iz jednog izvora je često ograničena, što zahteva kombinovanje podataka iz više izvora, čime se dodatno može povećati robusnost modela (Kamarudin et al., 2021). Na primer, u studiji (Wang et al., 2020) je ilustrovana upotreba podataka iz više izvora među kojima su satelitski snimci, klimatski podaci, mape zemljišta i istorijski zapisi o prinosu. Ovi podaci su kombinovani kako bi se izgradili napredni ML modeli za predviđanje prinosa ozime pšenice unutar sezone. Ovom studijom prevaziđena su ograničenja prethodnih pristupa zasnovanih isključivo na satelitskim snimcima i tradicionalnim statističkim metodama.

3. Teorijske osnove

U ovom poglavlju date su teorijske osnove za istraživanje sprovedeno u okviru ove doktorske disertacije. Izložena je teorijska osnova korišćenih metoda i pojmova važnih za razumevanje celokupne metodologije.

3.1 Uvod u veštačku inteligenciju

Veštačka inteligencija je oblast računarstva koja se bavi razvojem sistema sposobnih za obavljanje zadataka koji bi obično zahtevali ljudsku inteligenciju, uključujući složene misaone procese kao što su rasuđivanje, učenje i prilagođavanje (Goertzel & Pennachin, 2007).

3.1.1 Definicija i osnovna podela veštačke inteligencije

Iako se veštačka inteligencija može definisati na mnogo različitih načina, zajednički element svih definicija svodi se na sposobnost digitalnog računara ili robota pod kontrolom računara da obavlja zadatke koji su inače povezani sa inteligentnim bićima (Goertzel & Pennachin, 2007). Osnovna karakteristika veštačke inteligencije je da sistem prima ulazne podatke i na osnovu njih donosi određene odluke. Podaci koji služe kao ulaz u modele veštačke inteligencije mogu biti prikupljeni putem senzora, uređaja za snimanje ili pisanih zapisa, ali se uvek moraju pretvoriti u format koji je neophodan za računarsku obradu.

Zahvaljujući sposobnosti da analizira podatke i donosi odluke na osnovu toga, AI je našla primenu u mnogim oblastima, uključujući prepoznavanje glasa, autonomnu vožnju, pretraživače interneta i preciznu poljoprivredu.

Osnovna podela veštačke inteligencije napravljena je u zavisnosti od toga koliki je „stepen inteligencije“ sistema u poređenju sa čovekom (Russell & Norvig, 2016) i ona obuhvata dve grupe:

1. slabu (usku) veštačku inteligenciju (engl. *Narrow AI*),
2. jaku (opštu) veštačku inteligenciju (engl. *General AI*).

Savremene praktične aplikacije uglavnom koriste usku veštačku inteligenciju koja pokazuje sposobnost rešavanja pojedinačnih problema. Osnovni cilj ove vrste AI-a je da je nivo tačnosti ovih rešenja bude zadovoljavajući za praksu.

Za razliku od uskih sistema veštačke inteligencije dizajniranih za specifične zadatke, opšta veštačka inteligencija teži razvoju sistema sa sposobnošću razumevanja, učenja i primene znanja u širokom spektru domena. Takvi

sistemi teže postizanju opšte kognitivne fleksibilnosti slične ljudskoj (Goertzel, 2014).

Pored osnovne podele, u literaturi se može pronaći i podela veštačke inteligencije u odnosu na evolutivni napredak AI sistema u pogledu količine iskustva koje se koristi i dubine razumevanja sopstvenog delovanja. Prema klasifikaciji koju predlaže Hintze (2016), razlikujemo četiri različite kategorije:

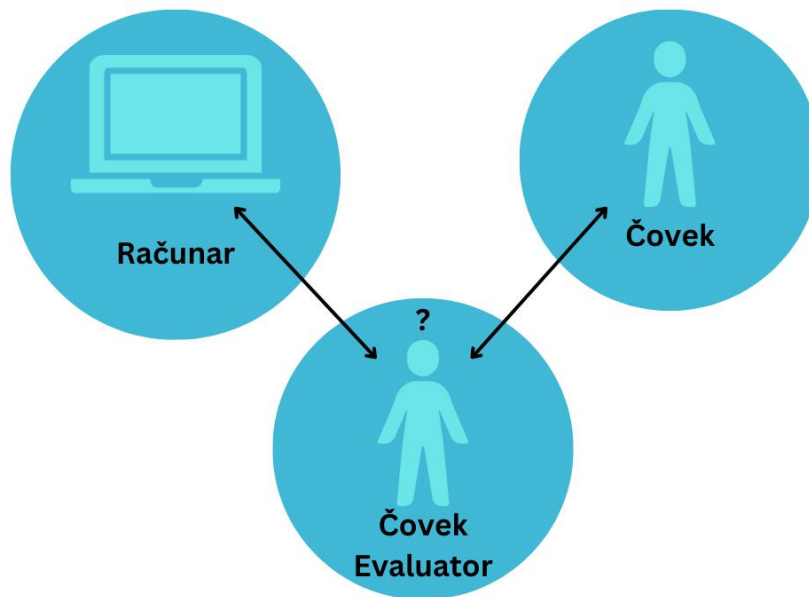
1. reaktivne mašine,
2. ograničeni memorijski sistemi,
3. sistemi sa teorijom uma,
4. samosvesni sistemi.

Reaktivne mašine rade isključivo na osnovu trenutnih ulaznih informacija, bez korišćenja pređašnjih iskustava. Ograničeni memorijski sistemi mogu koristiti istorijske podatke za donošenje odluka. Teorija uma podrazumeva sisteme koji mogu da „razumeju“ emocije i društvene interakcije, a samosvesna veštačka inteligencija je konceptualno razmatranje u kome mašine poseduju svest kao ljudi.

3.1.2 Istorijski razvoj veštačke inteligencije

Veštačka inteligencija je u proteklih sedam decenija doživela niz značajnih otkrića, tehnoloških inovacija, ali i neuspeha.

Teorijsku osnovu za veštačku inteligenciju postavio je Alan Turing još 1940-ih u svom radu *Računarska mašina i inteligencija*, u kome je uveo ideju o „univerzalnoj mašini“ sposobnoj da izvrši bilo koje matematičko izračunavanje (Turing, 1950). Pored toga, Turing je formulisao i čuveni Turingov test čiji je cilj da proceni da li mašina ima sposobnost da oponaša ljudsku inteligenciju. Ovaj test je i danas poznat kao jedan od centralnih koncepata u veštačkoj inteligenciji. Grafički prikaz testa nalazi se na slici 1.



Slika 1. Grafički prikaz Turingovog testa.

Termin „veštačka inteligencija“ je prvi put upotrebljen na *Dartmouth* konferenciji 1956. godine od strane Džona Makartija, Marvinina Minskog, Natanijela Ročestera i Kloda Šenona. Termin odražava njihovu viziju o mašinama koje mogu da oponašaju ljudski način razmišljanja i sposobnosti za rešavanje problema (McCarthy et al., 1955). U tom periodu pojavili su se rani AI programi, kao što je *Logic Theorist* (Newell et al., 1957), koji su pokazali da mašine mogu da rešavaju matematičke teoreme koristeći logičko rezonovanje.

Tokom 1950-ih i 1960-ih značajno se ulagalo u simboličku veštačku inteligenciju, takođe poznatu kao *good old-fashioned AI*. Razvijeni su sistemi koji su mogli da igraju šah, dokazuju teoreme, pa čak i razgovaraju na jednostavne načine. Primer za to je program ELIZA, chatbot koji je oponašao ljudski razgovor korišćenjem jednostavnih pravila i preformulisanjem korisničkog unosa (Weizenbaum, 1966).

Međutim, postojala su velika ograničenja ovih pristupa, posebno zbog oslanjanja na ručno definisana pravila i nemogućnosti skaliranja na složene situacije iz stvarnog sveta. To je dovelo do perioda koji je danas poznat kao „zima veštačke inteligencije“ (engl. *AI winter*). U periodu tokom 1970-ih i 1980-ih, napredak je stagnirao zbog mnogih faktora među kojima su nedovoljna računarska snaga, ograničeno finansiranje i neispunjena inicijalna očekivanja, što je dovelo do skepticizma u pogledu izvodljivosti AI (Crevier, 1993).

Veštačka inteligencija je doživela ponovni procvat kasnih 1990-ih i ranih 2000-ih, podstaknuta eksponencijalnim rastom računarske snage, pojavom velikih skupova podataka i novim dostignućima u mašinskom učenju. Posebno značajan bio je razvoj veštačkih neuronskih mreža i dubokog učenja, koje je, u stvari, podskup mašinskog učenja inspirisan strukturom i funkcijom ljudskog mozga. Duboke neuronske mreže su se u ovom periodu pokazale kao uspešne u zadacima kao što su prepoznavanje slika, obrada govora i razumevanje prirodnog jezika (LeCun et al., 2015).

Značajna dostignuća poput *Google*-ovog *AlphaGo*-a koji je pobedio svetskog šampiona u igri *Go* 2016. godine (Silver et al., 2016) i *IBM*-ovog *Deep Blue*-a koji je pobedio šahovskog šampiona Garija Kasparova 1997. godine (Campbell et al., 2002; Newborn, 2003), pokazala su kako veštačka inteligencija može da se bavi izazovima za koje se ranije verovalo da su isključivo namenjeni ljudima.

Zahvaljujući inovacijama velikih kompanija, veštačka inteligencija je u poslednje vreme izazvala više pažnje nego ikada ranije. Posebnu pažnju izazvao je OpenAI sa svojim revolucionarnim izumom – ChatGPT. ChatGPT je veliki jezički model (engl. *Large Language Model*, LLM) razvijen tako da omogućava efikasnu obradu i razumevanje prirodnog jezika (Nazir & Wang, 2023; Wang et al., 2023; Haque & Li, 2025). Primenom tehnika dubokog učenja, model je obučen na velikim količinama tekstualnih podataka, zbog čega može da prepoznaje obrasce u jeziku i generiše smislene, koherentne i kontekstualno prilagođene odgovore. ChatGPT poseduje veliki broj parametara koji mu omogućavaju da analizira korisnički unos i formuliše odgovore koji podsećaju na ljudsku komunikaciju. Zbog svoje sposobnosti da radi sa različitim temama i tipovima zadataka, postao je jedan od najčešće korišćenih alata za brojne primene. ChatGPT i slični modeli imaju velik potencijal da unaprede poslovne procese i promene način na koji ljudi komuniciraju sa digitalnim sistemima. Njihova primena doprinosi većoj efikasnosti, dostupnosti informacija i automatizaciji zadataka u različitim industrijama.

3.1.3 Značaj i primenljivost veštačke inteligencije

Danas je veštačka inteligencija prisutna u praktično svakom aspektu savremenog života – od alata za obradu prirodnog jezika i prilagođenih preporuka, preko autonomne vožnje, do medicinskih dijagnoza i predloga terapije. Sa rastom i razvojem veštačke inteligencije pokreću se rasprave o njenim etičkim pitanjima, društvenim uticajima i ograničenjima.

Uprkos navedenim izazovima, veštačka inteligencija je postala značajan faktor modernog društva. Njena vrednost dolazi iz njene sposobnosti da automatizuje složene procese, unapredi donošenje odluka i pruži efikasnija rešenja za globalne izazove. Veštačka inteligencija primenjena na velike količine podataka omogućava automatizaciju, što dalje vodi do smanjenja troškova i povećanja produktivnosti.

Jedna od najvažnijih primena veštačke inteligencije je njeno korišćenje u medicini. AI ima sposobnost da unapredi medicinsku dijagnostiku i tretman. Algoritmi veštačke inteligencije se koriste za analizu medicinskih slika, otkrivanje bolesti u ranoj fazi, personalizaciju terapija, optimizaciju resursa u zdravstvenim sistemima i drugo. Na primer, sistemi zasnovani na dubokim neuronskim mrežama pokazali su da mogu precizno da prepoznaju obrasce na radiološkim i patološkim slikama, olakšavajući rad medicinskim stručnjacima (McKinney et al., 2020; Esteva et al., 2017; Liu et al., 2019).

Veštačka inteligencija ima značajnu primenu i u poljoprivredi, gde otvara prostor za precizno upravljanje usevima, predviđanje prinosa i detekciju stresa kod biljaka. AI pomaže u optimizaciji upotrebe resursa kao što su voda (Gutiérrez et al., 2014) đubriva i pesticidi (Talaviya et al., 2020), čime doprinosi održivosti poljoprivrednih praksi (Araújo et al., 2021; de Oliveira & de S. e Silva, 2023). Veštačka inteligencija je korišćena za predviđanje prinosa pšenice na osnovu meteoroloških podataka i satelitskih slika, što je omogućilo brže i preciznije donošenje odluka u procesima uzgoja (Pettorelli et al., 2005; Xue & Su, 2017).

Industrijska proizvodnja takođe beleži značajan napredak zahvaljujući primeni AI-a u kontroli kvaliteta i automatizaciji proizvodnih linija (Tsang et al., 2022). Pametni roboti, koji funkcionišu na osnovu algoritama mašinskog učenja, doprinose kontroli kvaliteta i automatizaciji procesa. Oni, pored toga što obavljaju fizički zahtevne zadatke, uče iz podataka dobijenih iz novih situacija kako bi se bolje prilagodili budućim, neviđenim okolnostima i povećali efikasnost (Siciliano & Khatib, 2016; Kober et al., 2013).

AI se koristi i u transportu (Abduljabbar et al., 2019), gde igra ključnu ulogu u razvoju autonomnih vozila i optimizaciji logistike. Primer za to su algoritmi za planiranje ruta, predikciju gustine saobraćaja i bezbednosti na putevima koji smanjuju kašnjenja i povećavaju sigurnost. AI se takođe koristi u pametnim gradovima za upravljanje saobraćajem, optimizaciju energetske resursa i pružanje personalizovanih usluga građanima (Mahat et al., 2023; Ma et al., 2020; Chen et al., 2024; Dartmann et al., 2021).

Još jedna primena veštačke inteligencije je u obrazovanju. AI u ovoj oblasti doprinosi kreiranju personalizovanih kurikuluma, automatskom ocenjivanju i korišćenju interaktivnih nastavnih alata koji prilagođavaju sadržaj potrebama učenika. Ovi alati u kombinaciji sa *e-learning* platformama mogu da olakšaju učenje i povećaju njegovu efikasnost (Zhai et al., 2021; Halkiopoulos & Gkintoni, 2024; Firat, 2023).

3.2 Mašinsko učenje

Iako se pojmovi veštačke inteligencije, mašinskog učenja i dubokog učenja često pominju u sličnom kontekstu, ovi pojmovi nisu sinonimi. Mašinsko učenje je podskup veštačke inteligencije koji se odnosi na razvoj sistema koji mogu da uče i donose odluke na osnovu podataka bez eksplicitnog programiranja (Burkov, 2019), dok je duboko učenje podskup mašinskog učenja koji kao algoritme koristi višeslojne neuronske mreže (Goodfellow et al., 2016). Značenje veštačke inteligencije dato je u poglavlju 3.1.1, a na slici 2 se vidi međusobni odnos veštačke inteligencije, mašinskog i dubokog učenja.



Slika 2. Međusobni odnos veštačke inteligencije, mašinskog učenja i dubokog učenja.

Mašinsko učenje funkcioniše tako što obučava modele nad ulaznim podacima kako bi se poboljšao njihov učinak u donošenju odluka ili predviđanju. Proces uključuje tri glavne komponente:

1. ulazne podatke,
2. zadatak koji treba rešiti,
3. metriku evaluacije za procenu performansi modela.

Ulazni podaci se koriste za obuku modela, zadatak definiše cilj koji model treba da ostvari, a metrika evaluacije meri uspešnost modela u rešavanju tog zadatka (Jung, 2022).

Mašinsko učenje obuhvata širok spektar algoritama, od kojih je svaki namenjen rešavanju specifičnih vrsta problema. Modeli mašinskog učenja se na osnovu vrste algoritama koje koriste mogu podeliti u sledeće četiri grupe:

- Linearni modeli – spadaju u osnovne metode i oni se koriste za modelovanje jednostavnih statističkih odnosa između ulaznih i izlaznih promenljivih (James et al., 2013). Ovde spadaju linearna i logistička regresija.
- Modeli zasnovani na stablu – veoma su efikasni zbog svoje sposobnosti da prepoznaju složene obrasce i interakcije u podacima. U ovu grupu modela mašinskog učenja spadaju stabla odlučivanja, slučajne šume i metode povećanja gradijenta (Hastie et al., 2009).
- Bajesovski modeli – modeli mašinskog učenja koji koriste probabilistički pristup, oslanjajući se na Bajesovo zaključivanje za predviđanje integracijom prethodnog znanja i ažuriranjem uverenja kako novi podaci postanu dostupni (Murphy, 2012).
- Metode ansambla – modeli mašinskog učenja koji kombinuju predviđanja više modela kako bi poboljšali performanse, smanjili varijansu i postigli robusnije i pouzdanije rezultate (Zhou, 2012).

Ako posmatramo vrstu zadatka kojim se mašinsko učenje bavi, možemo ga podeliti u četiri glavne kategorije (Sarker, 2021; Lee i Shin, 2020):

- nadgledano učenje,
- nenadgledano učenje,
- polunadgledano učenje,
- učenje uz podsticaje.

Svaka od ovih kategorija nudi pristupe koji su prilagođeni specifičnim potrebama i ciljevima analize podataka.

3.2.1 Nadgledano učenje

Nadgledano učenje (engl. *Supervised Learning*) predstavlja grupu algoritama u kojima se model obučava nad skupom podataka koji su označeni, odnosno u

slučajevima kada su rezultati od interesa unapred poznati. Ovo omogućava modelu da „nauči“ pravilnosti i veze između ulaznih podataka i njihovih pripadajućih izlaza tokom procesa obuke (Cunningham et al., 2008; Bishop, 2006; Goodfellow et al., 2016).

Nadgledano učenje ima značajnu primenu u mnogim oblastima, uključujući procenu rizika, klasifikaciju slika, prediktivnu analitiku i otkrivanje prevara (Hastie et al., 2009). U zavisnosti od prirode izlazne promenljive, algoritmi nadgledanog učenja mogu se podeliti u dve osnovne grupe:

1. regresione algoritme,
2. algoritme klasifikacije.

Regresioni algoritmi se koriste za predviđanje kontinualnih numeričkih izlaza. U najčešće korišćene regresione metode spadaju linearna i polinomijalna regresija, koje modeliraju linearne i nelinearne odnose u podacima. Pored toga, postoje ansambl metode poput slučajne šume za regresiju (engl. *Random Forest Regression*, RFR) i gradijentno pojačavanje (engl. *Gradient Boosting*) koje pokazuju posebno dobre rezultate na složenijim skupovima podataka zahvaljujući svojoj sposobnosti da smanje varijansu i prilagode se nelinearnim zavisnostima. Za slučajeve kada tradicionalne metode nisu dovoljne, mogu se primeniti metod potpornih vektora (engl. *Support Vector Regression*, SVR) ili duboke neuronske mreže specijalizovane za regresione zadatke.

Algoritmi klasifikacije se primenjuju za kategorizaciju ulaznih podataka u diskretne klase. Osnovni model koji se može upotrebiti za binarnu i višeklasnu klasifikaciju je logistička regresija. Pored toga, često se koriste metode k-najbližih suseda (engl. *K-nearest Neighbors*, K-NN) i mašina potpornih vektora (engl. *Support Vector Machine*, SVM), koje koriste različite pristupe za klasifikaciju i određivanje granica između klasa. Često se koriste i probabilistički modeli poput Naivnog Bajesovog klasifikatora, koji efikasno rukuju velikim skupovima podataka korišćenjem statističkih metoda. Stabla odlučivanja i njihove ansambl varijante za klasifikaciju (slučajne šume i gradijentno pojačavanje) u širokoj su upotrebi zbog svoje interpretabilnosti i visoke tačnosti koju često postižu. Za kompleksnije zadatke kao što su prepoznavanje slika i obrada prirodnog jezika, primenjuju se duboke neuronske mreže, u koje spadaju i njihove konvolucione i rekurentne varijante.

U tabeli 4 su navedeni često korišćeni algoritmi nadgledanog učenja, kao i njihove prednosti i ograničenja.

Tabela 4. Pregled korišćenih algoritama mašinskog učenja sa njihovim osnovnim prednostima i ograničenjima.

Algoritam	Prednosti	Ograničenja
SVR	Može rešavati nelinearne probleme transformacijom u visokodimenzionalni prostor i robustan je na prilagođavanje sa dobrim odabirom hiperparametara.	Izbor odgovarajuće jezgrene funkcije može biti složen, metoda je računarski zahtevna za velike skupove podataka.
K-NN	Jednostavna implementacija, fleksibilnost za klasifikaciju i regresiju, bez pretpostavki o distribuciji podataka.	Performanse zavise od broja suseda (k) i metrike udaljenosti, metoda može biti spora kod velikih skupova podataka.
ANN	Sposobne su da modeliraju kompleksne i nelinearne odnose, imaju veliku prilagodljivost za različite primene.	Zahtevaju veliki skup podataka za efikasnu obuku, mogući problemi s prilagođavanjem i dugom obukom.
RFR	Robustan na prilagođavanje, postiže visoku tačnost u mnogim aplikacijama i dobro funkcioniše sa podacima koji imaju mnogo dimenzija.	Interpretacija modela može biti složena, veći računarski zahtevi za vrlo velika stabla ili skupove podataka.

U nastavku su ukratko objašnjeni korišćeni algoritmi, kao i njihovi primeri upotrebe u kontekstu rešavanja predikcije visine prinosa ulja i sličnih agronomskih problema.

3.2.1.1 Metod potpornih vektora za regresiju

Metod potpornih vektora se prvenstveno koristi za klasifikaciju identifikacijom optimalne prave ili hiperravni koja razdvaja različite klase tačaka podataka u zavisnosti od toga da li se nalaze na pozitivnoj ili negativnoj strani hiperravni.

Metod potpornih vektora za regresiju proširuje ovaj koncept na zadatke regresije. Široko se koristi u oblastima kao što je predviđanje prinosa useva (Cruz et al., 2017; Fujita et al., 2016; Sladojevic et al., 2016).

Jedna od ključnih prednosti SVR-a je njegova sposobnost da pojednostavi matematičku analizu transformacijom nelinearnih problema u ulaznom

prostoru u linearne probleme u visokodimenzionalnom prostoru obeležja (Fuentes et al., 2017).

3.2.1.2 Metod K-najbližih suseda

Metod K-najbližih suseda je nadgledani algoritam mašinskog učenja koji se koristi i za zadatke klasifikacije i za zadatke regresije.

U dosadašnjim istraživanjima je uspešno primenjivan za predviđanje prinosa useva (Sladojevic et al., 2016; Fuentes et al., 2017; Lu et al., 2017). K-NN algoritam pronalazi k najbližih uzoraka za nepoznate uzorke i predviđa ishod na osnovu prosečne vrednosti odgovora tih uzoraka.

Pre primene algoritma, potrebno je odrediti broj k . Optimalna vrednost za k se obično određuje putem unakrsne validacije, pri čemu se testiraju različite vrednosti k , obično između 2 i 30, kako bi se pronašlo najbolje rešenje.

3.2.1.3 Veštačka neuronska mreža

Veštačke neuronske mreže su uspešno primenjivane za predviđanje prinosa useva (Mohanty et al., 2016; Sladojevic et al., 2016; Kuo et al., 2023).

ANN se sastoji od ulaznog sloja, jednog ili više skrivenih slojeva i izlaznog sloja. U ovom kontekstu, za postizanje preciznih predviđanja prinosa useva koristi se neuronska mreža sa propagacijom unapred (engl. *feedforward*) sa algoritmom povratnog širenja greške (engl. *backpropagation*). Broj ulaznih neurona varira u zavisnosti od izabranih karakteristika, a primarni parametar podešavanja je broj skrivenih neurona, koji se prilagođava na osnovu skupa ulaznih karakteristika.

3.2.1.4 Slučajna šuma

Slučajna šuma (engl. *Random Forest*, RF) je nadgledani ansambl algoritam mašinskog učenja koji se koristi za zadatke klasifikacije i regresije.

Reč je o metodi ansambla što znači da se kombinuje više drugih modela, u ovom slučaju stabala odlučivanja, kako bi se poboljšala preciznost predviđanja. Za zadatke regresije, kao što je na primer predviđanje visine prinosa ulja, RF agregira predviđanja pojedinačnih stabala usrednjavanjem njihovih rezultata. Algoritam za ulazne parametre zahteva broj stabala u šumi i broj karakteristika koje se razmatraju pri svakom razdvajanju.

Zbog svoje interpretabilnosti i otpornosti na šum u podacima, RF je često primenjivan algoritam u poljoprivrednim istraživanjima i pokazao je visoku efikasnost u modelima predviđanja prinosa različitih useva (Sladojevic et al., 2016; Hagen et al., 2024; Chandel et al., 2021).

3.2.2 Nenadgledano učenje

Nenadgledano učenje (engl. *Unsupervised Learning*) podrazumeva analizu skupova podataka bez unapred definisanih oznaka, kako bi se identifikovali skriveni obrasci i odnosi u skupu podataka. U praksi je često skupo ili vremenski zahtevno obezbediti označene podatke, zbog čega se u tim slučajevima često koristi nenadgledano učenje.

Ova vrsta mašinskog učenja često se koristi u sistemima preporuka, gde na primer algoritmi na osnovu potrošačkih navika korisnika predlažu proizvode u prodavnicama. Takođe, često se koristi i u detekciji anomalija, analizi društvenih mreža i redukciji dimenzionalnosti.

Najčešće korišćene metode nenadgledanog učenja su:

- Klasterovanje – grupiše podatke na osnovu sličnosti. Algoritam *K-means* deli podatke u K klastera, pri čemu se svaka tačka dodeljuje klasteru čiji je centar najbliži posmatranoj tački (MacQueen, 1967). Postoji i hijerarhijsko klasterovanje koje se zasniva na drugačijem principu, odnosno kreiranju klastera spajanjem manjih ili podelom većih klastera (Hastie et al., 2009).
- Probabilističko grupisanje – izražava pripadnost podataka klasterima na osnovu verovatnoće, čime se bolje uzima u obzir nesigurnost u podacima (Bishop, 2006). Primer za ovu vrstu algoritama je model mešavine Gausovih distribucija (engl. *Gaussian Mixture Model*).
- Algoritmi povezivanja – identifikuju odnose između različitih stavki u skupu podataka. Često se koriste u analizi potrošačke korpe da bi se utvrdilo koji proizvodi se često kupuju zajedno kako bi se poboljšale preporuke kupcima (Agrawal et al., 1993).

3.2.3 Polu-nadgledano učenje

Polu-nadgledano učenje (engl. *Semi-Supervised Learning*) kombinuje elemente nadgledanog i nenadgledanog učenja. U ovom pristupu, model koristi malu količinu označenih podataka, koji usmeravaju proces učenja, dok većina podataka ostaje neoznačena (Chapelle et al., 2009; Zhu & Goldberg, 2009). Ovaj metod je naročito koristan u situacijama kada je označavanje

podataka finansijski ili vremenski zahtevno, ali postoji velika količina neoznačenih podataka koji mogu biti iskorišćeni.

Označeni podaci pomažu modelu da usmeri svoje učenje, dok neoznačeni podaci omogućavaju proširivanje skupa podataka i bolju generalizaciju modela. Na primer, model može prvo da koristi nenadgledano učenje da identifikuje klastere u podacima, a zatim primeni nadgledano učenje da označi ove klastere.

Generativne adverzijalne mreže (engl. *Generative Adversarial Networks*, GANs) mogu se primeniti u polu-nadgledanom učenju, pri čemu se dve neuronske mreže obučavaju u tandemu, jedna generiše podatke (generator), a druga ih ocenjuje (diskriminator) i na taj način poboljšava iskoristivost neoznačenih podataka (Goodfellow et al., 2016).

3.2.4 Učenje uz podsticaje

Učenje uz podsticaje uključuje algoritme u kojima agent uči kroz interakciju sa okruženjem putem povratnih informacija u obliku nagrada i kazni. Agent stupa u interakciju sa okruženjem, izvršava radnje za postizanje cilja i prima nagrade ili kazne na osnovu svog učinka. Vremenom, kroz niz pokušaja i grešaka, agent uči da optimizuje svoje ponašanje kako bi maksimizovao nagrade (Sutton & Barto, 2018).

Prednost ovog pristupa jeste što ne zahteva unapred označene podatke već koristi proces interakcije sa okruženjem kao osnovu za učenje. Ovaj pristup se obično koristi u razvoju igara i robotici za podučavanje mašina zadacima oponašajući ljudske procese donošenja odluka.

Primeri ovih algoritama su Q-učenje (engl. *Q-learning*) i duboka Q-mreža (engl. *Deep Q-Network*, DQN), koji se koriste za učenje donošenja odluka kroz interakciju sa okruženjem. U klasičnom Q-učenju koristi se tzv. Q-tablica – tabela u kojoj se za svaku kombinaciju stanja i akcije čuva brojčana vrednost, koja predstavlja očekivanu nagradu ako se u tom stanju preduzme ta akcija (Watkins & Dayan, 1992). Ako se agent nalazi u određenom stanju, on pomoću Q-tablice bira akciju koja ima najveću vrednost, tj. vodi ka najboljem mogućem ishodu. Međutim, ako postoji mnogo mogućih stanja i akcija (što je često u stvarnim problemima), Q-tablica postaje veoma velika i nepraktična za korišćenje. Zato se kod naprednijih algoritama, poput DQN-a, koristi neuronska mreža koja uči da predviđa te vrednosti umesto da ih čuva eksplicitno u tabeli (Mnih et al., 2015). Na taj način omogućeno je efikasnije učenje u složenim i visoko-dimenzionalnim okruženjima.

3.3 Veštačke neuronske mreže

Veštačke neuronske mreže (engl. *Artificial Neural Networks*, ANN) predstavljaju jedan od najčešće korišćenih pristupa mašinskog učenja. One predstavljaju parametrizovanu reprezentaciju koja može poslužiti za aproksimaciju drugih funkcija. Kao i u slučaju drugih metoda učenja, pronalaženje odgovarajućih parametara se vrši matematičkom optimizacijom nekog kriterijuma kvaliteta aproksimacije i može biti računarski veoma zahtevno.

Osnovna gradivna jedinica neuronskih mreža je neuron koji je inspirisan funkcionisanjem bioloških neurona (slika 3). Prvi matematički model neurona razvili su Mekalok (engl. *McCulloch*) i Pits (engl. *Pitts*), čime su postavili temelje za savremeni razvoj veštačkih neuronskih mreža (McCulloch et al., 1943).

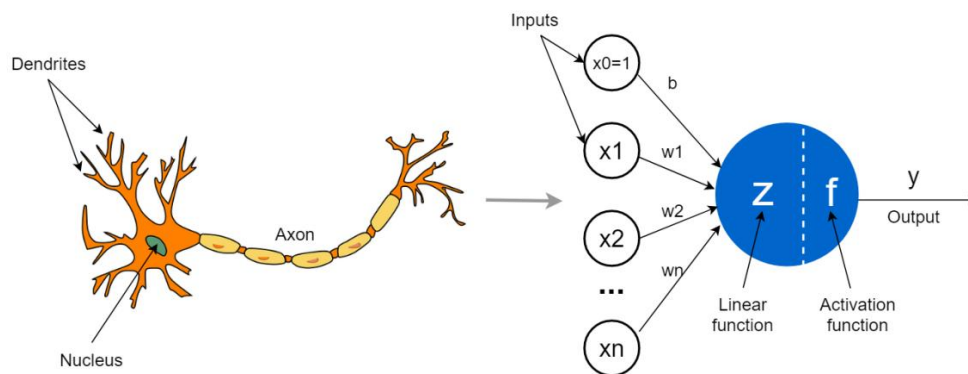
Matematički model neurona opisuje jedinicu za izračunavanje koja na osnovu ulaznih vrednosti, težinskih faktora (engl. *weights*) i dodatnog parametra (engl. *bias*) generiše izlaznu vrednost. Težinski faktori se inicijalno nasumično dodeljuju i prilagođavaju tokom procesa učenja mreže kako bi se minimizovala greška modela.

Jedan od osnovnih modela veštačkog neurona je perceptron (Rosenblatt, 1958), koji je 1958. godine predstavio Frenk Rozenblat (engl. *Frank Rosenblatt*). Perceptron funkcioniše kao jednostavan linearni klasifikator. Izračunava linearni zbir ulaznih vrednosti pomnoženih odgovarajućim težinama, uz dodatak praga b , prema formuli:

$$z = \sum (w_i * x_i) + b$$

Izlaz perceptrona se određuje primenom aktivacione funkcije $\varphi(z)$ (najčešće stepenasta funkcija) na zbir. Ako je zbir veći ili jednak nuli, izlaz je 1, u suprotnom je 0:

$$y = \varphi(z) = \{1, \text{ ako je } z \geq 0; 0, \text{ ako je } z < 0\}$$



Slika 3. Biološki neuron i matematički model veštačkog neurona (Ahmed i Lindblad, 2022).

Perceptron se koristi za rešavanje problema koji su linearno razdvojivi. Iako ima ograničenja, kao što je nemogućnost rešavanja nelinearno razdvojivih problema (npr. XOR problema), postao je osnova za razvoj složenijih modela neuronskih mreža, poput višeslojnih perceptrona (engl. *multilayer perceptron*), koji koriste slojeve skrivenih neurona i nelinearne aktivacione funkcije za obradu kompleksnijih podataka.

Višeslojni perceptron je neuronska mreža koja sadrži dodatne slojeve neurona između ulaznog i izlaznog sloja. Te dodatne slojeve nazivamo skrivenim slojevima. Zbog postojanja ovih slojeva, model je u stanju da „uči“ i rešava kompleksnije probleme.

Obuka višeslojne neuronske mreže se odvija u nekoliko koraka:

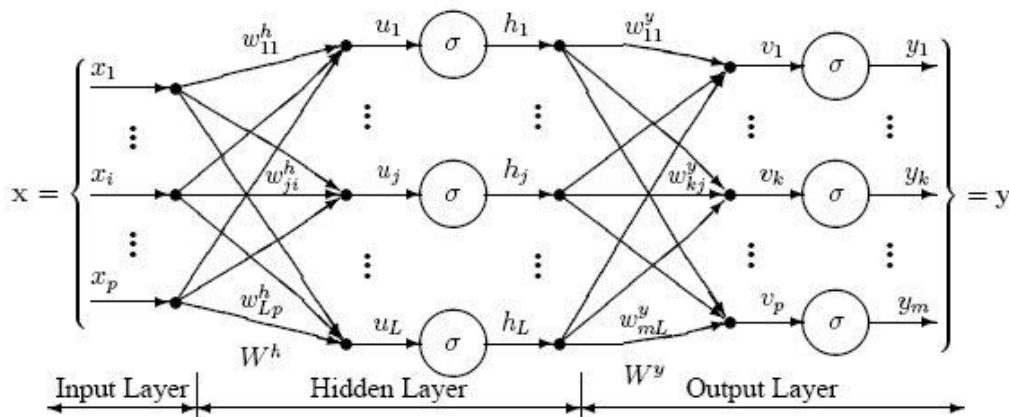
1. Prelaz unapred (engl. *forward pass*) – U ovom koraku izračunavaju se izlazne vrednosti neuronske mreže za zadati ulaz. Korak počinje inicijalizacijom težinskih parametara, a nakon toga podaci prolaze kroz neuronsku mrežu. Svaki neuron u svakom sloju višeslojnog perceptrona računa svoj izlaz primenom aktivacione funkcije na sumu težinskih parametara. Podaci se kreću u jednom smeru u toku ovog procesa. Na kraju, rezultat je izlazna vrednost mreže za zadate ulazne podatke.
2. Računanje greške (engl. *loss calculation*) – Nakon prelaza unapred, računa se greška, odnosno koliko se dobijeni izlaz razlikuje od željene vrednosti. Koriste se različite funkcije greške u zavisnosti od problema koji se rešava.
3. Propagacija greške unazad (engl. *backpropagation*) – U ovom koraku, greška se propagira unazad kroz mrežu kako bi se ažurirale inicijalne težine korišćenjem algoritma gradijentnog spusta (engl. *gradient descent*) koji omogućava minimiziranje greške.

Prilikom obučavanja neuronske mreže bira se funkcija gubitka koja predstavlja meru razlike između predviđenih i stvarnih vrednosti i koristi se za procenu uspešnosti modela tokom obučavanja. Kod binarne klasifikacije se najčešće koristi binarna unakrsna entropija (engl. *binary cross-entropy*) koja je posebno pogodna kada podaci nisu klasno uravnoteženi.

Funkcija gubitka se minimizuje podešavanjem težinskih parametara mreže. Za njihovo ažuriranje mogu se koristiti različiti optimizacioni algoritmi, među kojima je ADAM koji je poznat po svojoj efikasnosti i adaptivnoj sposobnosti za učenje (Kingma & Ba, 2017).

Kada se ovi koraci ponove za svaki ulaz iz skupa podataka za obučavanje, završava se jedna epoha obuke. Proces obučavanja se nastavlja sve dok se ne dostigne unapred određen broj epoha ili dok kada greška ne padne na prihvatljiv nivo.

Proces obučavanja modela može se pratiti pomoću vizuelnog prikaza krivih tačnosti (engl. *accuracy*) i gubitka (engl. *loss*) tokom obučavanja. Prikazane krive omogućavaju razumevanje ponašanja modela tokom obuke, posebno u pogledu stabilnosti konvergencije i sposobnosti generalizacije.



Slika 4. Obuka višeslojne neuronske mreže (Rao et al., 2015).

3.4 Duboko učenje

3.4.1 Uvod u duboko učenje

Termin duboko učenje označava disciplinu u okviru veštačke inteligencije i mašinskog učenja koja se zasniva na korišćenju dubokih neuronskih mreža i principa kompletnog učenja (engl. *end-to-end learning*) na osnovu izvornih podataka (Halkiopoulous & Gkintoni, 2024). Ideja kod dubokog učenja jeste da

se neuronske mreže sastoje od više slojeva, gde svaki sloj omogućava učenje hijerarhijskih reprezentacija podataka.

Najznačajnije primene dubokog učenja ostvarene su u oblasti računarske vizije. To obuhvata rešavanje sledećih problema:

1. klasifikacije slika,
2. detekcije i lokalizacije objekata,
3. semantičke segmentacije.

Klasifikacija ili kategorizacija slika smatra se jednim od osnovnih problema u računarskoj viziji. Ovaj problem je polazna tačka za mnoge druge probleme u računarskoj viziji. Tradicionalni pristup problemu klasifikacije slika zasnivao se na ručnom izdvajanju karakteristika (fičera) slike na osnovu kojih bi se obučavao klasifikator (Hastie et al., 2009; Ahmed & Lindblad, 2022). Iako su se neuronske mreže i ranije koristile (Rao et al., 2015), veliki napredak u ovoj oblasti nastupa uvođenjem konvolucionih neuronskih mreža. Kod detekcije i lokalizacije objekata, cilj je naći različite objekte na slici i odrediti njihove okvirne oblasti. Segmentacija ide korak dalje i vrši klasifikaciju svakog piksela slike. Arhitekture koje se koriste za pomenute zadatke zasnivaju se na konvolucionim neuronskim mrežama za klasifikaciju uz određena proširenja.

Pored primene dubokog učenja u računarskoj viziji, duboke neuronske mreže su pokazale primenu i u kreativnijim zadacima što se vidi na primeru prenošenja umetničkog stila sa jedne slike na drugu (engl. *style transfer*) (Gatys et al., 2015) ili menjanja sadržaja metodom „haluciniranja” ili „dubokog sna” (engl. *deep dream*) (Mordvintsev et al., 2015). Takođe, pokazalo se da duboke neuronske mreže mogu da se koriste i za kreiranje novog sadržaja generisanjem slika. Ovaj tip neuronskih mreža predstavljaju generativne mreže.

Još jedna vrsta neuronskih mreža su rekurentne neuronske mreže (engl. *Recurrent Neural Network*, RNN) koje se bave obradom sekvencijalnih podataka, poput rečenica prirodnog jezika i vremenskih serija. One modeluju zavisnost između elemenata sekvence, gde broj bitnih prethodnih elemenata nije fiksno određen. RNN obrađuju elemente sekvence korak po korak, koristeći skriveno stanje koje akumulira informacije iz prethodnih koraka (Rumelhart et al., 1986). Parametri mreže određuju kako se skriveno stanje ažurira i kako se generiše izlaz, omogućavajući obradu dinamičkih i zavisnih podataka (Kanellopoulos & Wilkinson, 1997). Primeri upotrebe ove vrste neuronskih mreža su u problemima prepoznavanja govora (engl. *speech recognition*) (Graves, 2013; Hannun et al., 2014), sinteze govora (engl. *text-to-speech*), generisanja tekstualnih opisa za slike (u kombinaciji sa CNN) itd.

Kod dubokog učenja uz podsticaje postoji agent koji opaža trenutno stanje okruženja i preduzima akcije zbog kojih dobija nagrade predstavljene numeričkim vrednostima u odnosu na koje prelazi u novo stanje (Mnih et al., 2015). Duboko pojačano učenje je zaslužno za najnovija dostignuća u domenu igranja logičkih i strateških igara. Za razliku od konvencionalnog pristupa ovom problemu, duboko pojačano učenje se zasniva na učenju, umesto na ugrađivanju pravila. Prva demonstracija primene ove vrste dubokog učenja bila je kada je program *AlphaGo* uspeo da pobedi aktuelnog šampiona u igri *Go* (Silver et al., 2016).

3.4.2 Razlike između dubokog i mašinskog učenja

Važno je naglasiti razlike između dubokog i mašinskog učenja. Mašinsko učenje koristi algoritme za analizu, učenje i donošenje odluka na osnovu podataka (Bishop, 2006; Hastie, 2009). Međutim, konvencionalne tehnike mašinskog učenja imaju ograničenu sposobnost obrade podataka u njihovom originalnom obliku. Zbog toga su ove metode zahtevale značajno razumevanje i stručnost za predstavljanje, odnosno odabir i pripremu odgovarajućih karakteristika iz podataka.

Duboko učenje predstavlja podskup mašinskog učenja koji omogućava automatsko izdvajanje, analizu i razumevanje dobijenih informacija na osnovu podataka (Goodfellow et al., 2016). Rezultati dobijeni dubokim učenjem su u mnogim situacijama bili bolji od konvencionalnog pristupa mašinskom učenju (Arik et al., 2017). Duboke neuronske mreže se sastoje od više skrivenih slojeva koji omogućavaju da sistem uči o složenom odnosu između ulaza i izlaza.

Prednost dubokog učenja u odnosu na mašinsko učenje je da ne zahteva ručno ekstrahovane ili ručno izrađene karakteristike koje se očekuju u mašinskom učenju. Dakle, duboko učenje automatski izdvaja karakteristike iz neobrađenih podataka, obrađuje ih i donosi odluku na osnovu toga (Karpathy & Fei-Fei, 2015).

3.5 Konvolucione neuronske mreže

Konvoluciona neuronska mreža je tip neuronske mreže koja sama uči karakteristike putem optimizacije filtera (kernela). Ova vrsta neuronske mreže se primenjuje za obradu mnogo različitih tipova podataka uključujući tekst, slike i audio zapis (LeCun et al., 1998; LeCun et al., 2015; Goodfellow et al., 2016).

Ovaj tip neuronske mreže dobio je ime po konvoluciji, operatoru koji se primenjuje u obradi slike i signala. U okviru ove mreže se koriste konvolucionni filteri koji se koriste za izoštravanje i zamućenje slika, kao i za detektovanje ivica, pomoću kojih se detektuju objekti na slici. CNN su napravljene tako da prevazilaze neke nedostatke klasičnih, potpuno povezanih neuronskih mreža, gde se kao ulazni podaci koriste slike. S obzirom na to da se pikseli slika mapiraju na ulazne neurone, zbog broja ulaznih neurona, postaje teško ili praktično nemoguće obučavanje potpuno povezanih neuronskih mreža. Ako je reč o slikama u boji, dimenzije ulaznih neurona se dodatno uvećavaju, jer je potrebno uzeti u obzir i 3 kanala kojima su predstavljene osnovne boje (RGB).

CNN se sastoje iz više slojeva za otkrivanje bitnih osobina ulaznih podataka. Zbog toga se konvolucionni filteri primenjuju na sliku kako bi se izvukle korisne karakteristike i kreirale njihove mape karakteristika (engl. *feature maps*). Na taj način se, pored otkrivanja značajnih karakteristika, vrši i redukcija rezolucije. Nedostatak CNN je potreba za značajnim hardverskim resursima u odnosu na potpuno povezane neuronske mreže.

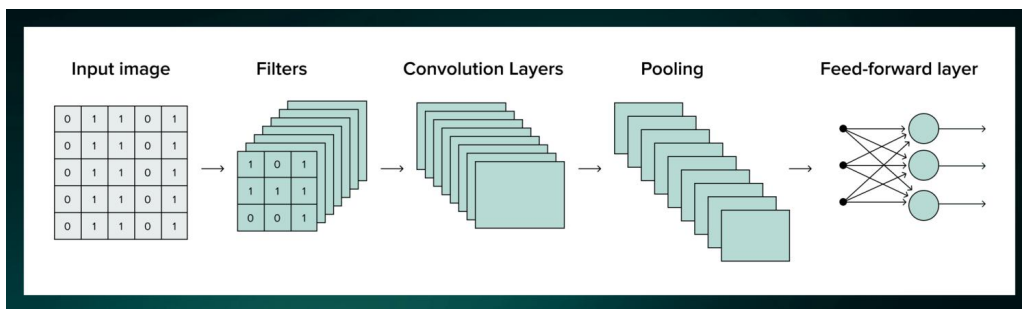
Konvolucione neuronske mreže, poznate su i kao veštačke neuronske mreže sa prostornom invarijantnošću, jer koriste arhitekturu deljenih težina – konvolucione filtere (jezgra) koji se kreću duž ulaznih podataka. Ova karakteristika omogućava mreži da prepozna obrascе nezavisno od njihove prostorne pozicije, pri čemu se kao rezultat dobijaju mape obeležja (Muñoz, 2014).

3.5.1 Arhitektura konvolucione neuronske mreže

Iako postoje različite varijacije arhitektura konvolucionih neuronskih mreža, standardne konvolucione neuronske mreže uvek se sastoje od tri obavezna sloja gradivnih blokova koji je formiraju, a to su:

1. konvolucionni sloj,
2. sloj/slojevi sažimanja,
3. potpuno povezani sloj.

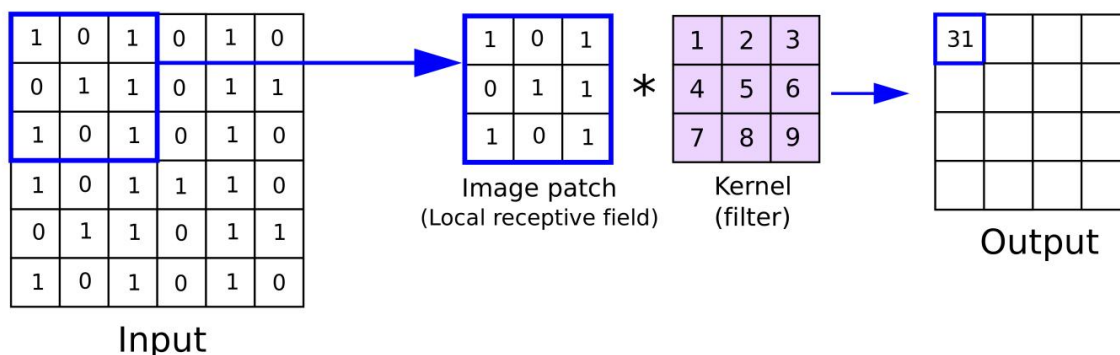
Arhitektura konvolucione neuronske mreže na kojoj se vide njena osnovna tri sloja prikazana je na slici 5.



Slika 5. Arhitektura konvolucione neuronske mreže (Singh, 2024).

3.5.1.1 Konvolucionni sloj

Konvolucionni sloj baziran je na operatoru konvolucije. Konvolucija u konvolucionim neuronskim mrežama (slika 6) funkcioniše primenom filtera (kernela), koji je u stvari mala matrica, obično dimenzija 3×3 ili 5×5 , koja se pomera preko ulaznih podataka. Ovaj proces naziva se pomicanje (engl. *stride*), pri čemu se kernel translira unapred definisanim koracima duž slike ili podataka. Tokom ovog postupka, vrednosti filtera se množe sa odgovarajućim vrednostima u ulaznoj matrici, a zatim se rezultati sabiraju kako bi se dobila nova vrednost u izlaznoj matrici. Krajnji rezultat ove operacije je mapa karakteristika, koja predstavlja transformisanu verziju ulaznih podataka sa izdvojenim ključnim karakteristikama.



Slika 6. Konvolucija u konvolucionim neuronskim mrežama (Nagib et al., 2022).

Konvolucija u konvolucionim neuronskim mrežama može se klasifikovati prema različitim kriterijumima, u zavisnosti od načina primene i svrhe.

Prema broju ulaznih i izlaznih kanala, razlikuju se:

1. jednostruka (engl. *single-channel*),
2. višekanalna (engl. *multi-channel*) konvolucija.

Kod jednostruke konvolucije, filter se primenjuje na jednokanalne slike (npr. crno-bele), dok se kod višekanalne konvolucije filteri koriste na slikama sa

više kanala (npr. RGB slike), gde se rezultati sumiraju preko svih kanala. Prema načinu pomeranja kernela, konvolucija može biti:

1. standardna (sa zadatim korakom),
2. konvolucija sa proširenim poljem receptora (engl. *dilated convolution*).

Kod konvolucije sa proširenim poljem receptora, elementi kernela se raspoređuju sa prazninama kako bi obuhvatili širi kontekst bez povećanja broja parametara.

Prema veličini izlazne mape karakteristika, razlikujemo:

1. *valid*,
2. *same*,
3. *full* konvoluciju.

Valid konvolucija ne koristi dopunu (engl. *padding*), što smanjuje dimenzije izlazne slike, same konvolucija zadržava iste dimenzije kao ulaz, dok full konvolucija koristi dodatni padding kako bi obuhvatila i piksele na ivicama slike. Ove različite vrste konvolucije omogućavaju prilagođavanje mreže različitim zadacima u obradi slika i drugih vrsta podataka.

U konvolucionim slojevima se koriste i aktivacione funkcije koje uvode nelinearnost kako bi mreža mogla da uči složene obrasce. ReLU (engl. *Rectified Linear Unit*) je najčešće korišćena funkcija jer propušta pozitivne vrednosti, dok negativne postavlja na nulu, čime sprečava problem nestajanja gradijenta i ubrzava obuku. Pored ReLU funkcije, često se koristi i *Leaky ReLU* koja rešava problem „mrtvih neurona“ dozvoljavajući male negativne vrednosti. Postoje i Sigmoid i Tanh koje se retko koriste zbog problema sa zasićenjem gradijenta, ali se mogu primeniti u specifičnim slučajevima.

3.5.1.2 Sloj sažimanja

Motivacija iza korišćenja sloja sažimanja je smanjenje veličine reprezentacije uz zadržavanje samo važnijih karakteristika. Utvrđeno je da slojevi sažimanja funkcionišu veoma dobro, ali osnovni razlozi za to nisu u potpunosti shvaćeni.

Najčešći tip koji se koristi je tehnika sažimanja maksimumom (engl. *max pooling*) koja se zasniva na tome da se određeni region piksela zamenjuje maksimalnom vrednošću piksela u datom regionu. Na ovaj način najizraženije informacije se propagiraju dalje kroz neuronsku mrežu.

Nekada se u arhitekturama konvolucionih neuronskih mreža koristila i tehnika sažimanja srednjom vrednošću (engl. *average pooling*), međutim ona više nije toliko zastupljena u savremenim arhitekturama.

3.5.1.3 Potpuno povezan sloj

Potpuno povezan sloj (engl. *fully connected layer*) je sloj neuronske mreže u kome je svaki neuron povezan sa svim neuronima u prethodnom i sledećem sloju.

U konvolucionim neuronskim mrežama, potpuno povezani slojevi se pojavljuju pri kraju arhitekture i igraju ključnu ulogu u kombinovanju ekstrahovanih karakteristika za donošenje odluka. Nakon što slojevi konvolucije i sažimanja izdvoje karakteristike iz slike, potpuno povezani slojevi obrađuju ove karakteristike da bi izvršili klasifikaciju ili regresiju. U zavisnosti od zadatka, ovaj sloj može koristiti i aktivacione funkcije, poput Softmax funkcije koja se koristi za klasifikaciju. Ova funkcija koristi se da bi se numerički izlazi mreže pretvorili u verovatnoće.

3.5.2 Najčešće korišćene vrste konvolucionih neuronskih mreža

Od uvođenja LeNet-5 – prve razvijene arhitekture konvolucionih neuronskih mreža, razvijeno je više različitih arhitektura, od kojih je svaka unapređivala prethodne dizajne u smislu dubine, efikasnosti i performansi. Ove arhitekture su evoluirale kroz značajne inovacije, baveći se izazovima kao što su prilagođavanje, nestajući gradijenti i računarska neefikasnost. Sa porastom dubokog učenja, CNN-ovi su postali osnovna komponenta u različitim aplikacijama, uključujući primene u medicinskim slikama, autonomnu vožnju i obradu prirodnog jezika. Istraživači su kontinuirano optimizovali dizajn CNN-a uvođenjem novih mrežnih struktura, aktivacionih funkcija i tehnika obuke kako bi poboljšali efikasnost i tačnost učenja.

U nastavku su nabrojane najvažnije CNN arhitekture, njihov istorijski značaj i ključni doprinos unapređenju dubokog učenja:

- LeNet-5 (1998) – LeNet-5 su razvili Yann LeCun et al. (1998). Ova arhitektura je jedna od najranijih CNN-a dizajniranih za prepoznavanje rukom pisanih cifara (LeCun et al., 1998). LeNet-5 se sastoji od dva konvoluciona sloja, dva sloja sažimanja i potpuno povezana sloja.
- AlexNet (2012) – AlexNet su predstavili Krizhevsky et al. (2012). Ova arhitektura je tada osvojila *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC), nakon čega je značajno povećano interesovanje za duboko učenje. Korišćena je ReLU aktivacija za bržu obuku i rano odustajanje (engl. *dropout*) kako bi se sprečilo prilagođavanje.
- VGG (2014) – Arhitekture *Visual Geometry Group* (VGG) (Simonyan & Zisserman, 2015) uvele su duboke mreže sa malim konvolucionim filterima (dimenzija 3x3). VGG-16 i VGG-19 su dve varijacije ove

arhitekture koje su pokazale da dubina poboljšava sposobnost učenja, ali i usporava vreme obuke. VGG je jedna od sporijih u poređenju sa drugim savremenim arhitekturama konvolucionih neuronskih mreža.

- GoogLeNet (2014) – GoogLeNet (Szegedy et al., 2015) je predstavio Inception modul, koji paralelno koristi više konvolucionih filtera različitih veličina. GoogLeNet, iako sačinjena od 22 sloja (27 slojeva ako se računaju i slojevi sažimanja), ima svega oko 4 miliona težina za obučavanje, za razliku od AlexNet arhitekture koja ima oko 60 miliona parametara za obučavanje. Ova arhitektura je poboljšala efikasnost uz smanjenje potrebe za hardverom kroz smanjenje dimenzionalnosti.
- ResNet (2016) – Rezidualne mreže (ResNet), koje su predložili He et al., (2016), uvele su rezidualno učenje, koje omogućava obuku izuzetno dubokih mreža (npr. ResNet-50, ResNet-101, ResNet-152). Novitet ove arhitekture predstavljaju rezidualni blokovi (engl. *residual blocks*) koji omogućuju obučavanje veoma „dubokih“ neuronskih mreža jer rešavaju problem nestajućih gradijenata (engl. *vanishing gradient*).
- DenseNet (2017) – DenseNet (Huang et al., 2017) povezuje svaki sloj sa svakim sledećim, promovišući ponovnu upotrebu funkcija i smanjujući broj parametara u poređenju sa ResNet-om, poboljšavajući efikasnost.
- EfficientNet (2019) – EfficientNet (Tan & Le, 2019) uvodi složenu metodu skaliranja koja optimalno balansira dubinu, širinu i rezoluciju, značajno poboljšavajući tačnost uz zadržavanje efikasnosti.

U tabeli 5 navedene su prednosti često korišćenih arhitektura u kontekstu malih skupova podataka.

Tabela 5. Prednosti često korišćenih arhitektura u kontekstu malih skupova podataka.

Arhitektura	Prednosti u kontekstu malih skupova podataka
VGG16/VGG19	Jednostavna implementacija i visoka interpretabilnost, dobre performanse na manjim skupovima podataka primenom transfernog učenja, stabilna i robustna arhitektura.
DenseNet121	Efikasno koristi informacije iz prethodnih slojeva, poboljšavajući generalizaciju na malim skupovima podataka, ima manji broj parametara u poređenju sa dubokim mrežama poput VGG.
InceptionV3	Dobro se prilagođava karakteristikama, što je korisno sa malim skupovima podataka, ima manje parametara, što može ubrzati obuku i smanjiti preprilagođavanje.
MobileNet	Efikasna u pogledu broja parametara, što omogućava bržu obuku nad malim skupovima podataka.

3.6 Transferno učenje

Transferno učenje je pristup u mašinskom učenju koji poboljšava proces učenja novog zadatka korišćenjem znanja stečenog iz prethodno naučenog zadatka. Za razliku od tradicionalnog mašinskog učenja, gde se modeli obučavaju izolovano za svaki konkretan problem, transferno učenje omogućava modelima da primene prethodno stečeno znanje na nove probleme. Ovo je posebno značajno za zadatke u kojima su podaci za obuku modela oskudni ili ih je teško pribaviti.

Osnovna ideja koja stoji iza transfernog učenja je da mnogi zadaci mašinskog učenja dele zajedničke strukture i obrasce. Umesto da se svaki put počne od nule, ideja je da modeli ponovo koriste naučene obrasce, što može da dovede do brže konvergencije, veće tačnosti i smanjenja potrebne količine označenih podataka.

Transferno učenje se široko koristi u dubokom učenju, posebno u kombinaciji sa konvolucionim neuronskim mrežama. Razlog za to je što postoje unapred obučeni modeli koji su obučeni na velikim skupovima podataka koji se mogu prilagoditi odnosno „fino podesiti“ (engl. *fine-tuning*) za nove, specifične primene. *ImageNet* je jedan od najpoznatijih skupova podataka koji se koriste za transferno učenje i on obuhvata više od 14 miliona označenih slika (Deng et al., 2009).

Efikasnost transfernog učenja zavisi od stepena sličnosti između izvornih i ciljnih zadataka, količine dostupnih označenih podataka i prenosivosti naučenih funkcija. Kada su izvorni i ciljni zadaci slični, model može efikasno preneti stečeno znanje, čime se smanjuje potreba za velikim količinama podataka i resursa za ponovno obučavanje (Pan & Yang, 2010). Nasuprot tome, ako postoji značajna razlika između domena primene, performanse modela mogu se pogoršati, što može zahtevati dodatna, fina podešavanja ili prilagođavanja specifičnih slojeva neuronske mreže (Goodfellow et al., 2016).

Količina označenih podataka u ciljnom domenu takođe igra važnu ulogu. Transferno učenje je posebno korisno kada je skup podataka ograničen, jer omogućava modelima da iskoriste znanje iz velikih, bogato označenih skupova podataka kako bi poboljšali performanse na manjim skupovima. Međutim, u slučajevima kada su podaci ciljnog domena obimni i dobro označeni, prednosti transfernog učenja mogu biti manje izražene.

Treći ključni faktor je prenosivost naučenih funkcija, odnosno koliko dobro prethodno naučeni obrasci mogu da se primene na novi zadatak. U

slučajevima kada model uči opšte karakteristike podataka koje su relevantne za više domena, prenosivost je visoka. Na primer, duboki konvolucionni neuronski modeli trenirani na velikim skupovima podataka poput *ImageNet*-a često se koriste kao osnova za različite vizuelne zadatke, jer prenose opšte vizuelne karakteristike (Yosinski et al., 2014).

Postoje različite strategije za implementaciju transfernog učenja, uključujući:

- Izdvajanje karakteristika (engl. *feature extraction*) – Korišćenje prethodno obučenih slojeva modela za izdvajanje relevantnih karakteristika i njihovu primenu na novi klasifikator.
- Fino podešavanje (engl. *fine tuning*) – Odmrzavanje nekih slojeva prethodno obučenog modela i njihovo ponovno obučavanje na novom skupu podataka kako bi se prilagodili zahtevima specifičnim za zadatak.
- Prilagođavanje domena (engl. *domain adaptation*) – Prilagođavanje modela obučenog na jednom domenu da radi dobro u drugom, povezanom domenu.

3.7 Tehnike pripreme podataka

3.7.1 Eksploratorna analiza podataka

Eksploratorna analiza podataka (engl. *Exploratory Data Analysis*, EDA) je skup metoda za ispitivanje podataka koja koristi njihove glavne karakteristike i vizuelizuje ih. EDA se fokusira na proveru pretpostavki potrebnih za prilagođavanje modela i testiranje hipoteza, rukovanje vrednostima koje nedostaju i pravljenje transformisanih varijabli ako je potrebno. EDA nam pruža vizuelni i numerički prikaz podataka opisujući skupove podataka, broj redova/kolona, podatke koji nedostaju, tipove podataka, itd. Takođe, sadrži metode za čišćenje oštećenih podataka i rukovanje podacima koji nedostaju, nevažecim tipovima podataka i netačnim vrednostima (Tukey, 1977; Sahoo et al., 2019).

3.7.2 Tehnike augmentacije podataka

Augmentacija slika u skupu podataka se koristi da bi se rešio problem preprilagođavanja. Tehnike augmentacije se mogu podeliti na osnovu faze u kojoj se primenjuju na:

- *Offline* tehnike augmentacije
- *Online* tehnike augmentacije

Offline tehnike podrazumevaju da se augmentacija primenjuje pre same faze obučavanja i da se rezultati skladište u fizičkoj memoriji. Ovaj pristup se

najčešće koristi kada se obučava nad manjim skupom podataka (Tang et al., 2020).

Online pristup se primenjuje kada su skupovi podataka veliki i kada postoji ograničenje u pogledu hardverskih resursa. U tom slučaju, augmentacija se vrši tokom treninga nad manjim delovima podataka (engl. *mini batch size*), što omogućava efikasniji rad sa većim bazama podataka (Talukdar et al., 2018).

4. Metodologija

U ovom poglavlju predstavljene su dve ključne oblasti istraživanja. U prvom delu opisana je metodologija za detekciju stresa usled suše kod biljaka suncokreta, dok je u drugom opisana metodologija za predviđanje visine prinosa ulja kod biljaka suncokreta.

4.1 Metodologija za detekciju stresa usled suše

Eksperimentalni dizajn za detekciju stresa usled suše kod suncokreta predstavljen je u pet faza. Prva faza predstavlja opis eksperimentalne postavke, odnosno opis kontrolisanih uslova pod kojima je izvršen eksperiment. Druga faza obuhvata prikupljanje podataka i opis njihovih karakteristika. Nakon toga sledi priprema za analizu i ulaz u modele. U ovoj fazi akcenat je na primeni različitih metoda obogaćivanja podataka u cilju poboljšanja efikasnosti korišćenog modela dubokog učenja. Sledeća faza bavi se kriterijumima za odabir modela dubokog učenja, a obuka izabranih modela i njihova evaluacija su prikazane na kraju, u poslednjoj fazi.

Grafički prikaz predložene metodologije dat je u obliku procesnog toka (engl. *pipeline*) na slici A.1.

4.1.1 Opis eksperimentalne postavke

Eksperimenti koji su korišćeni za dobijanje podataka za detekciju stresa usled suše u biljkama suncokreta izvedeni su u komori za rast dimenzije 2,60x4,18x3,08m. Komora se nalazi u Institutu za ratarstvo i povrtarstvo, u Rimskim Šančevima, kod Novog Sada, Srbija (45° 19' 44,7204" N, 19° 49' 39,29916" E).

Eksperiment je sproveden u kontrolisanim uslovima kako bi se ispitala reakcija inbred linija suncokreta na uslove suše. Inbred linije predstavljaju genetski uniformne i visoko homozigotne linije koje nastaju višestrukim samoprašivanjem i selekcijom u cilju zadržavanja poželjnih genetičkih osobina (Poehlman & Sleper, 1995; Acquaah, 2009). Korišćeni su podaci o 12 inbred linija suncokreta, koje su odabrane na osnovu otpornosti na sušu, produktivnosti i otpornosti na gljivične bolesti. Izgled klima komore predstavljen je na slici 7.



Slika 7. Eksperimentalna postavka - biljke suncokreta u klima komori.

Inbred linije su uzgajane u tri kontrastna statusa vlažnosti zemljišta izražena u gravimetrijskom sadržaju vode (engl. *gravimetric water content*, GWC) koji su korišćeni za simuliranje različitih nivoa stresa usled suše:

- dobro navodnjeno zemljište – 70% GWC (kontrola),
- stres suše – 50% GWC i
- ekstremni stres suše – 42% GWC.

U eksperimentu su korišćeni rizotroni koji su omogućili kontinuirano posmatranje bez ometanja biljke ili tla (Klepper & Kaspar, 1994).

Postavljanje neinvazivne platforme za fenotipizaciju rizotрона u objektima za uzgoj kontrolisanog okruženja urađeno je u skladu sa Nagel et al. (2012) i modifikovano da odgovara korenovom sistemu suncokreta i uslovima rasta suncokreta.

Eksperiment je sproveden pod veštačkim osvetljenjem (*Philips HPI-T 1000V/543*, metal-halogenid, hladno bela, 82100 Lm) na temperaturi od 23 $\pm$ 2°C, 16/8h svetlo/mrak, i oko 70% relativne vlažnosti vazduha. Unutrašnje dimenzije rizotрона korišćenih u eksperimentima su bile 29x59x1,5cm. Rizotroni su postavljeni pod uglom od 45° kako bi se povećala vidljiva dužina korena na donjoj, proвидноj strani rizotрона. Svaki rizotron je napunjen supstratom *Klasmann Deilmann Potground P* (zamrznut kroz crni treset), a

GWC supstrata je određen korišćenjem alata *Moisture Analyzer DBS 60-3 Kern & Sohn GmbH*.

Seme suncokreta je pre setve stavljeno na vlažan filter papir 24 sata na temperaturi od 23°C, u mraku. Za svaki rizotron postavljeno je po jedno seme koje je već počelo da klija (dužine korenka klice 1-2 mm), na dubini od 1 cm. Površina tla je zatim prekrivena slojem perlita od 1 cm da se smanji gubitak vode.

Jedan rizotron se smatrao jednom replikom. Korišćeno je deset replika po inbred liniji i tretmanu i one su nasumično postavljene. Biljke su uzgajane u rizotronima sve dok primarni koren nije dostigao dno rizotrona sa GWC od 70%. Izgled biljke u rizotronu prikazan je na slici 8.



Slika 8. Biljka suncokreta u rizotronu.

4.1.2 Opis skupa podataka

Podaci koji su korišćeni za detekciju stresa usled suše kod biljaka suncokreta su prikupljeni iz opisanog eksperimenta tokom 2023. godine u obliku slika snimljenih kamerom i skenerom. Eksperiment i dobijeni podaci su omogućili

praćenje fizičkih promena na delovima biljaka suncokreta u odnosu na različite nivoe sadržaja vode u zemljištu.

Skup podataka koji je korišćen kao ulaz u modele mašinskog učenja sastoji se od slika delova biljaka suncokreta dobijenih na dva različita načina.

Prvi način prikupljanja podataka bilo je fotografisanje korena biljaka iz rizotrona koje predstavlja neinvazivno prikupljanje informacija jer se koren ne vadi iz zemlje. Slike su fotografisane kroz providne panele rizotrona kako bi se analizirala i prikazala morfologija korena pri različitim nivoima sadržaja vode u zemljištu.

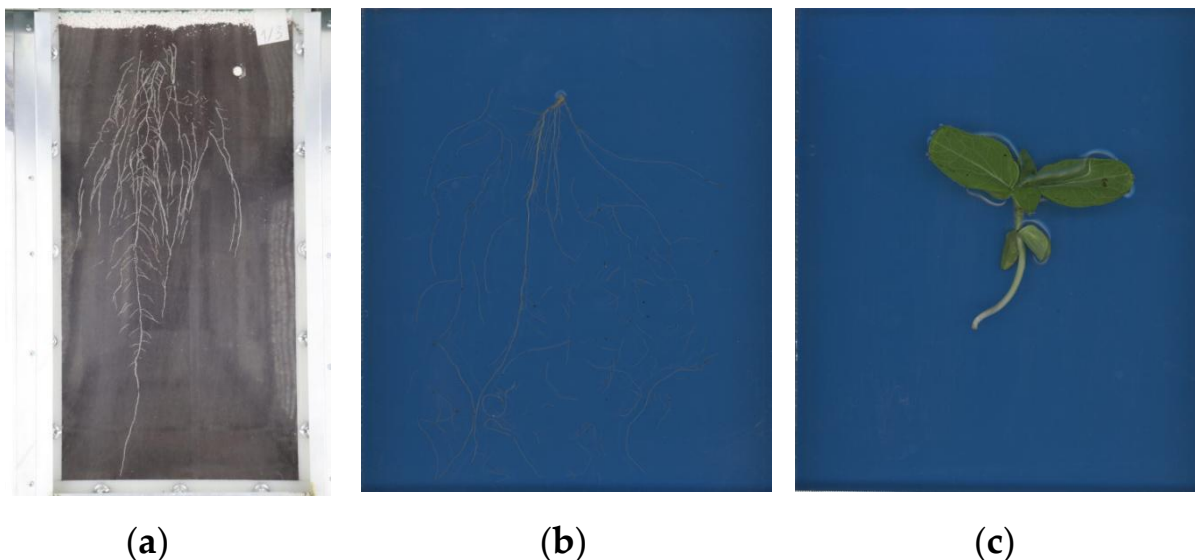
Fotografije korena u rizotronima snimljene su u kutiji sa veštačkim svetlom i kamerom (*Canon EOS 850D – 24,1 MP*) postavljenom na fiksnoj udaljenosti, 14. dana eksperimenta. Rezolucija fotografija bila je 4000x6000 piksela, a fotografije su sačuvane u JPEG formatu. Konačni broj slika koji je prikupljen na opisan način bio je 209.

Drugi način prikupljanja podataka podrazumevao je skeniranje korena suncokreta (nakon vađenja iz rizotrona) i njihovih nadzemnih organa kako bi se dobio širi pogled na efekte stresa usled suše.

Četrnaestog dana eksperimenta, nakon što su biljke fotografisane, korenov sistem je uklonjen iz rizotrona, zemljište je isprano i korenovi su fino očišćeni od ostataka zemljišta. Posle pranja, svaki koren je stavljen u providnu tacnu sa količinom vode dovoljnom da se koren u nju potopi. Korenovi su postavljeni na skener (dimenzije 25x20cm) tako da se izbegne preklapanje i pojava dva ili više korena koji se lepe jedan za drugi. U slučaju velikog korenovog sistema, korenje je isečeno na nekoliko delova kako bi se izbegli prethodno navedeni problemi. Parametri u *WinRhizo Pro* softveru (*Regent Instruments Canada Inc., Québec, KC, Kanada*), među kojima su veličina ležišta (20x25 cm), izbor praga sive (180-205) i vertikalna i horizontalna rezolucija (400 dpi), postavljeni su pre skeniranja da bi se isključili ostaci i podesili nivoi sive za vidljivost.

Pripremljeni korenski sistem i nadzemni delovi biljaka su zatim skenirani i izmereni pomoću *WinRhizo Pro* softvera. Konačni broj slika koji je prikupljen na opisan način bio je 375.

Finalni skup podataka predstavlja kombinaciju podataka koji su prikupljeni na dva opisana načina, te je ukupan broj slika bio 584. Reprezentativni primeri slika koje su korišćene u eksperimentu mogu se videti na slici 9.



Slika 9. Reprezentativni primeri skupa podataka: (a) fotografija korena suncokreta iz rizotrona; (b) skenirani korenov sistem; (c) skenirani nadzemni deo biljke.

4.1.3 Obrada i obogaćivanje podataka

Konačni skup podataka podeljen je na trening, validacioni i testni skup. Trening skup je obuhvatao 80% slika, pri čemu je 20% tog skupa korišćeno za validaciju. Preostalih 20% ukupnog skupa podataka činilo je testni skup, pri čemu je obezbeđeno da testni skup ne sadrži podatke koji su korišćeni u trening skupu, odnosno da nisu uključene slike iste inbred linije koje su već bile deo trening skupa. Ovi podaci su zatim korišćeni za obuku i evaluaciju modela dubokog učenja sa ciljem predviđanja stresa usled suše.

Pošto su skupovi podataka koji su korišćeni u istraživanju relativno mali u poređenju sa onima koji su obično potrebni za postizanje najboljih rezultata u dubokom učenju, korišćenje transfernog učenja je važan korak u poboljšanju performansi i robusnosti modela. Međutim, pretrenirani modeli koji se koriste u transfernom učenju su trenirani nad opštim skupovima podataka, a ne nad domen specifičnim kao što su slike biljaka suncokreta pod različitim nivoima vodnog stresa.

Standardne tehnike augmentacije obuhvataju dodavanje različitih varijanti slika okretanjem, rotacijom i zumiranjem postojećih. Pored toga, koristi se i ciljano uvećanje podataka specifično za karakteristike konkretnog skupa podataka koje može dodatno da poboljša rezultate.

Takođe, korišćenje slika u sivim tonovima, uz augmentaciju podataka, može otvoriti put ka boljim rezultatima i dati nove uvide u istraživanje.

U eksperimentima koji su sprovedeni u ovom radu, korišćene su offline i online tehnike augmentacije podataka. Detalji korišćene *offline* augmentacije navedeni su u potpoglavlju 4.1.3.3. Pored toga, korišćene su i *online* tehnike augmentacije da bi se ojačala otpornost modela na promene u orijentaciji slike, svetlosti, razmeri i deformaciji.

Za augmentaciju trening skupa i normalizaciju svih skupova podataka korišćene su instance klase ImageDataGenerator iz biblioteke Keras (Chollet, 2015). Parametri augmentacije su odabrani na osnovu preliminarnih eksperimenata i literature (Patra & Sahoo, 2024), pri čemu rezultati ukazuju da veći opsezi dovode do distorzija kod biljnih slika, dok manji ne obezbeđuju dovoljnu raznolikost. Odabrani su sledeći parametri:

- horizontalno okretanje,
- opseg rotacije: ± 20 stepeni,
- opseg pomeranja širine: $\pm 20\%$,
- opseg pomeranja visine: $\pm 20\%$,
- opseg zuma: $\pm 20\%$,
- opseg osvetljenosti: 0,1–0,9.

4.1.3.1 Klasna neuravnoteženost i prilagođavanje eksperimentalnog dizajna

Skup podataka iz istraživanja je bio klasno neuravnotežen, što predstavlja jedan od najčešćih problema kod klasifikacionih modela (Shorten et al., 2021). Najmanje je bilo fotografija biljaka koje su bile pod najvećim sušnim stresom, a najviše je bilo fotografija dobro zalivenih biljaka, jer su one zbog svoje veličine često zahtevale više kadrova, što je produkovalo više slika.

Do disbalansa je došlo i jer značajan broj biljaka izloženih sušnom stresu od 42% nije uopšte uspeo da raste, što sugeriše da je nivo suše bio prevelik. Da bismo rešili ovaj problem i pojednostavili proces klasifikacije, prilagodili smo naš eksperimentalni dizajn. Umesto da kategorišemo biljke u tri različite grupe na osnovu tačnih procenata GWC-a (42%, 50% i 70%), grupisali smo ih u dve široke kategorije klasifikacije, i to: biljke pod stresom od suše i biljke bez stresa (dobro zalivene).

Ovo prilagođavanje ne samo da odražava praktične rezultate nivoa stresa na rast biljaka, već i pomaže u upravljanju neuravnoteženosti klasa unutar skupa podataka konsolidacijom kategorija stresa.

Međutim, čak i nakon prilagođavanja eksperimentalnog dizajna, skup podataka je ostao neuravnotežen. Bilo je 228 fotografija biljaka koje su iskusile stres usled suše (42% i 50% GWC) i 356 fotografija dobro zalivenih biljaka (70% GWC). Zbog ovoga su primenjene strategije augmentacije

podataka na CNN modelu koji je postigao najbolje rezultate nakon inicijalnog eksperimenta.

4.1.3.2 Generisanje dodatnih uzoraka deljenjem slika korenovog sistema

Drugi pristup za povećanje raznolikosti i reprezentativnosti skupa podataka bilo je proširenje skupa podataka dodavanjem novih slika koje su kreirane na osnovu postojećih (Das et al., 2021; An et al., 2019; Virnodkar et al., 2020). Glavni fokus bio je na tome da se nadomesti nedostatak slika koje prikazuju delimične korenove pod stresom usled suše. Ovo ograničenje je nastalo jer korenovi pod stresom obično postaju manji i često se u potpunosti uklapaju u jedan okvir slike. Nasuprot tome, delimični prikazi korenova su bili prisutni u skupu podataka za biljke koje nisu bile pod sušnim stresom jer su njihovi korenovi bili veći.

Kroz augmentaciju, kreirane su dodatne slike koje bolje reprezentuju ove specifične scenarije, čime je poboljšan kvalitet i raznolikost skupa podataka. Kreiranjem novih slika na osnovu postojećih zatvoren je jaz i simuliran širi spektar scenarija stresa usled suše. Nove slike su kreirane deljenjem 13 nasumično odabranih originalnih slika korenova iz trening skupa koje pripadaju grupi biljaka pod stresom od suše, horizontalno i vertikalno.

U grupi sa dobro zalivenim biljkama, bilo je 356 slika, od čega 66 slika koje su sadržale delove biljaka (nekompletne biljke), što čini približno 18,5% ukupnih slika sa dobrim zalivanjem. Zbog održavanja proporcionalnosti između klasa, odabrali smo približno isti procenat slika iz grupe biljaka pod stresom usled suše, odnosno 13 slika.

4.1.4 Odabir modela dubokog učenja

Najpopularniji model dubokog učenja za rad s dvodimenzionalnim ulaznim podacima, kao što su slike, je konvoluciona neuronska mreža. CNN arhitekture su posebno osmišljene da efikasno procesiraju podatke organizovane u obliku mreže, kao što su pikseli u slikama, zbog čega su prvi izbor u zadacima prepoznavanja i analize slika.

Primena dubokog učenja u agronomiji dovela je do promena u procesu fenotipizacije biljaka na osnovu slika (Singh et al., 2018). Neke od arhitektura dubokog učenja koje su korišćene za fenotipizaciju biljaka su DCNN (Azimi et al., 2021), RCNN, ResNet, VGGNet, ResNeXt, LeNet, ZFNet (Rojanarungruengporn & Pumrin, 2021), GoogLeNet, AlexNet (Rojanarungruengporn & Pumrin, 2021; Chandel et al., 2022), SqueezeNet (Khalifani et al., 2022), GAN (Călin et al., 2023). Ove arhitekture su se pokazale kao dobar izbor u širokom spektru zadataka fenotipizacije biljaka,

kao što su identifikacija biljaka na osnovu obrazaca lisnih vena (Malik et al., 2022), brojanje listova (Ghosh et al., 2023), brojanje resica kod kukuruza (Çetin et al., 2021), brojanje stabljika biljke i određivanje njene širine (Lužaić et al., 2023), segmentacija metlica (Badshah et al., 2024), lokalizacija korenova i izdanaka i detekcija karakteristika (Pound et al., 2017), detekcija cvetanja (Curran, 1987) i prepoznavanje vrsta biljaka (Kurtulmuş, 2021).

U našem istraživanju primenljeno je pet CNN arhitektura, i to: VGG16, VGG19, Inception V3, DenseNet121 i MobileNet.

Svi ovi modeli su prethodno obučeni na velikom skupu podataka *ImageNet*, koji sadrži više od 14 miliona označenih slika. Obuka na *ImageNet*-u omogućava modelima usvajanje osnovnih reprezentacija vizuelnih karakteristika koje se mogu prilagoditi različitim zadacima, uključujući i zadatke specifične za određene domene, poput agronomije. Zbog svoje složene i duboke arhitekture, ovi modeli pružaju solidnu osnovu za prilagođavanje, dok njihova dizajnerska rešenja omogućavaju postizanje visokih performansi u prepoznavanju obrazaca na slikama (Ali et al., 2024).

Izbor pomenutih CNN arhitektura bio je rezultat detaljne analize i razmatranja nekoliko ključnih faktora.

Kao prvi faktor, u obzir su uzete specifičnosti dizajna svake arhitekture. Na primer, VGG arhitekture su poznate po jednostavnosti strukture i upotrebi slojeva iste veličine filtera, što olakšava interpretaciju rezultata. S druge strane, Inception koristi koncept ugnježđenih slojeva, omogućavajući efikasnije prepoznavanje složenih obrazaca. DenseNet se ističe po tome što povezuje svaki sloj s prethodnim slojevima, čime se smanjuje problem gubitka gradijenta, dok MobileNet cilja na aplikacije koje zahtevaju manju potrošnju resursa, što je ključno za primenu u realnim sistemima sa ograničenim hardverskim resursima.

Takođe smo uzeli u obzir performanse ovih arhitektura na *ImageNet* skupu podataka, koji se često koristi kao standard za procenu kvaliteta CNN modela. Ovi modeli su pokazali dosledno visoke rezultate na zadacima klasifikacije slika, što ih čini pogodnim za prilagođavanje specifičnim potrebama našeg istraživanja (tabela 6).

Tabela 6. Performanse CNN arhitektura na *ImageNet* skupu (Keras, 2024).

Arhitektura	Broj parametara (milioni)	Veličina modela (MB)	Top-1 tačnost na skupu <i>ImageNet</i> (%)
VGG16	138.4	528	71.3

VGG19	143.7	549	71.3
InceptionV3	23.9	92	77.9
DenseNet121	8.1	33	75
MobileNet	4.3	16	70.4

Efikasnost u pogledu računarskih resursa bila je još jedan značajan faktor prilikom izbora CNN modela. Vodio se računa o tome da odabrane arhitekture mogu biti fino podešene unutar ograničenja naše dostupne infrastrukture. Ovo je posebno važno s obzirom na činjenicu da zadaci obuke i testiranja dubokih neuronskih mreža često zahtevaju značajne resurse, kako u pogledu procesorske snage, tako i u pogledu memorije.

Pored tehničkih razloga, izbor CNN modela zasnovan je i na prethodnim uspešnim primenama ovih arhitektura u kontekstu agronomije (Chandel et al., 2021; Barradas et al., 2021). Naime, navedena istraživanja su pokazala da su izabrani modeli efikasni u različitim zadacima, kao što su detekcija bolesti biljaka, klasifikacija useva i procena prinosa, što je dodatno potvrdilo njihov potencijal za predviđanje stresa izazvanog sušom kod biljaka suncokreta.

Na kraju, još jedan razlog za izbor ovih arhitektura bila je njihova dostupnost preko *TensorFlow Hub*-a (TensorFlow Developers, 2022). Ova platforma omogućava jednostavan pristup unapred obučanim modelima, čime se značajno smanjuje vreme potrebno za početnu fazu razvoja modela.

Korišćenje unapred obučanih arhitektura omogućilo je da fokus u istraživanju bude na finom podešavanju modela za specifičan zadatak umesto da se troše resursi na obuku modela od nule. Upoređene su performanse različitih modela i identifikovana je najefikasnija arhitektura CNN-a, koja je kasnije korišćena kao bazni model za razvoj modela za detekciju stresa usled suše.

4.1.5 Priprema i obuka modela dubokog učenja

Okvir konvolucione neuronske mreže za detekciju stresa usled suše na osnovu slika suncokreta koji je predložen u ovom istraživanju sastoji se od šest slojeva. Svi slojevi su prilagođeni za konkretan problem istraživanja, odnosno efikasno prepoznavanje sušnih stanja kod suncokreta.

Prvi sloj je ulazni sloj, čija je glavna uloga da prihvati slike kao ulazne podatke. Slike su prethodno pripremljene za ulaz tako da imaju iste dimenzije i format koji odgovara ulaznim zahtevima modela, čime je omogućeno dosledno i efikasno obučavanje mreže.

Drugi sloj se zasniva na unapred obučenom CNN modelu, koji je treniran na *ImageNet* skupu podataka i predstavlja osnovni model u arhitekturi. Osnovni model funkcioniše kao model za izdvajanje karakteristika, pružajući bogat skup dubokih karakteristika ključnih za razlikovanje stresa usled suše od drugih potencijalnih varijacija na slikama. Strategija korišćenja unapred obučениh modela značajno smanjuje računarsko opterećenje i omogućava brzo prilagođavanje specifičnostima našeg domena.

Nakon osnovnog modela nalazi se treći sloj, odnosno sloj globalnog prosečnog sažimanja (engl. *global average pooling*), koji je uveden kako bi se smanjila dimenzionalnost podataka i računarska složenost. Ovaj sloj agregira informacije iz izlaza osnovnog modela, pretvarajući ih u kompaktnu reprezentaciju koja zadržava ključne karakteristike potrebne za dalju obradu.

Četvrti sloj je potpuno povezan gusti sloj (engl. *fully connected dense layer*), koji koristi ReLU kao aktivacionu funkciju. Četvrti sloj uvodi nelinearnost u model, omogućavajući mu da se bolje prilagodi kompleksnim obrascima u podacima. Dodavanje ovog sloja pomaže u unapređivanju sposobnosti modela da prepozna suptilne razlike između biljaka pod stresom i onih koje su bile dobro zalivene.

Peti sloj je sloj nasumičnog isključivanja neurona (engl. *dropout layer*) sa stopom isključivanja od 0,5, strateški postavljen da smanji preprilagođavanje (Shin et al., 2016). Ovaj sloj nasumično izostavlja određeni procenat čvorova tokom obuke, čime model postaje robusniji i manje zavisn od specifičnih uzoraka u podacima. Uvođenjem ovog sloja povećava se sposobnost generalizacije modela, što je ključno za njegovu primenu na novim, neviđenim podacima.

Poslednji, šesti sloj je završni, gusti sloj sa jednim neuronom. Ovaj neuron koristi sigmoidalnu aktivacionu funkciju koja mapira izlaz modela na vrednosti između 0 i 1, što predstavlja verovatnoću pripadnosti klasi biljke u odnosu na stres usled suše. Na ovaj način, model proizvodi jasan i interpretabilan rezultat.

Model je sastavljen koristeći Adam optimizator, a kao funkcija gubitka korišćena je binarna unakrsna entropija.

Efikasnost predloženog modela procenjena je na osnovu standardnih metrika koje se obično koriste u binarnoj klasifikaciji sa neuravnoteženim skupovima podataka. Korišćene metrike – tačnost, preciznost, odziv i F1-skor obezbedile su sveobuhvatan uvid u performanse modela, što je detaljno opisano u poglavlju 4.1.6. Pored toga, korišćeni su grafikoni kao što su ROC krive i

AUC vrednosti za evaluaciju sposobnosti modela da pravilno klasifikuje biljke u kontekstu sušnog stresa.

Svih pet CNN modela više puta je obučeno sa različitim ručno podešenim parametrima kako bi se pronašli optimalni parametri i poboljšale performanse. Ovo je obavljeno i za originalne i za skupove podataka u sivim tonovima da bi se odredile optimalne stope učenja, veličine serija i broj epoha.

U okviru pripreme podataka pre ulaza u model, veličina slika je promenjena na 224×224 , a vrednosti piksela su podeljene sa 255 da bi se skalirale u opseg od 0 do 1. Ovaj korak pojednostavljuje podatke, ubrzavajući učenje i konvergenciju neuronskih mreža.

Testirane su veličine serije od 16, 32 i 64 i pokazalo se da 32 daje najbolje početne rezultate za sve modele. Ova veličina obezbeđuje dobru računarsku efikasnost bez iscrpljivanja sistemske memorije. Veličina serije od 64 može dovesti do ograničenja memorije i manje stabilnih procena gradijenta, dok je veličina serije od 16 bila manje efikasna i usporila je proces obuke. Utvrđeno je da je VGG16 zahtevao nižu stopu učenja od 0,00001 i duži period obuke od 50 epoha. Nasuprot tome, VGG19, DenseNet121, InceptionV3 i MobileNet su uspešno obučavani tokom 30 epoha sa višom stopom učenja od 0,0001 (tabela 7).

Korišćene su slike u sivim tonovima da bi se utvrdilo da li su informacije o boji neophodne za tačno otkrivanje stresa usled suše kod suncokreta. Ovaj pristup je zasnovan na prethodnim istraživanjima u vezi sa primenom veštačke inteligencije u poljoprivredi u kojima su pojednostavljene metode snimanja, kao što su slike u sivim tonovima, efikasno korišćene za otkrivanje i kategorizaciju stresa biljaka (Liu et al., 2017; Ramcharan et al., 2017).

Tabela 7. Optimalan broj epoha i *learning rate* za korišćene modele.

CNN Model	Broj epoha	<i>Learning rate</i>
VGG16	50	0.00001
VGG19	30	0.0001
DenseNet121	30	0.0001
Inception V3	30	0.0001
MobileNet	30	0.0001

4.1.6 Metode evaluacije

Metrike korišćene za evaluaciju performansi i poređenje različitih CNN arhitektura u zadatku detekcije stresa usled suše bile su: tačnost (*Accuracy*), preciznost (*Precision*), odziv (*Recall*) i F1 mera (*F1 Score*).

Tačnost je metrika koja kvantifikuje ukupne performanse klasifikatora (jednačina 1).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (\text{Jednačina 1})$$

Preciznost meri proporciju istinitih pozitivnih predviđanja u odnosu na sva pozitivna predviđanja koja je napravio klasifikator (jednačina 2).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (\text{Jednačina 2})$$

Procenat istinitih pozitivnih ishoda među svim stvarnim pozitivnim rezultatima meri se odzivom (jednačina 3).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (\text{Jednačina 3})$$

F1 mera je metrika koja predstavlja harmonijsku sredinu između preciznosti i odziva i koristi se kao pokazatelj ukupne efikasnosti modela u smislu pozitivnih predviđanja klasa (jednačina 4).

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (\text{Jednačina 4})$$

U prethodnim jednačinama su upotrebljene sledeće oznake:

- TP (*True Positives*) – broj slučajeva koji su tačno predviđeni kao pozitivni.
- FP (*False Positives*) – broj slučajeva koji su pogrešno predviđeni kao pozitivni kada su zapravo negativni.
- TN (*True Negatives*) – broj slučajeva koji su tačno predviđeni kao negativni.
- FN (*False Negatives*) – broj slučajeva koji su pogrešno predviđeni kao negativni kada su zapravo pozitivni.

4.2 Metodologija za predviđanje visine prinosa ulja

Eksperimentalni dizajn za predviđanje visine prinosa ulja kod suncokreta je predstavljen u nekoliko faza. Prva sadrži opis eksperimentalne postavke, odnosno kontrolisanih uslova pod kojima je izvršen eksperiment. Druga faza obuhvata postupke prikupljanja podataka i njihove karakteristike. Nakon toga sledi eksploratorna analiza podataka, u okviru koje su podaci objašnjeni i vizuelno prikazani. Najčešće korišćeni modeli mašinskog i dubokog učenja su prikazani u narednoj fazi, dok su obuka izabranih modela i njihova evaluacija prikazani u poslednjoj fazi.

4.2.1 Opis eksperimentalne postavke

Eksperimenti postavljeni u cilju razvoja modela za predviđanje visine prinosa suncokretovog ulja obuhvatali su poljska istraživanja na različitim lokalitetima Panonskog regiona. Prikupljeni podaci su sadržali informacije o prinosu semena, sadržaju ulja u semenu, kao i detaljnijim meteorološkim i zemljišnim parametrima specifičnim za svaku lokaciju.

Eksperimenti su sprovedeni na Panonskom području koje je poznato po povoljnim uslovima za gajenje suncokreta i gde je suncokret značajna uljana biljna vrsta. Korišćeno je 1250 hibrida suncokreta, razvijenih i selektovanih u Institutu za ratarstvo i povrtarstvo u Novom Sadu, naučnoj instituciji koja se dugi niz godina bavi oplemenjivanjem biljaka. Hibridi su testirani u različitim agroekološkim uslovima, čime je obezbeđena reprezentativnost skupa podataka za analizu uticaja mikroklimatskih i zemljišnih faktora na prinos i kvalitet semena suncokreta.

Ispitivanje hibrida je obavljeno na četiri različite lokacije: Novi Sad (NS), Subotica (SU), Kikinda (KI) i Vršac (VŠ). Lokacije su odabrane zbog ekoloških razlika u pogledu tipa zemljišta, padavina i temperature, čime su omogućeni različiti specifični mikroklimatski uslovi za uzgoj suncokreta. Pored toga, navedene lokacije su već korišćene za ispitivanje novih hibrida. Podaci su prikupljeni za dve vegetacione sezone: 2018. i 2019. godinu.

Meteorološki podaci, u koje spadaju temperatura i količina padavina, su zabeleženi iz Republičkog hidrometeorološkog zavoda Srbije za svaku od lokacija. Temperatura je praćena na dnevnom nivou, kako bi se omogućila analiza temperaturnih oscilacija, a podaci o padavinama su beleženi kako bi se procenili uslovi vlažnosti tla tokom ključnih faza vegetacije. Takođe, prikupljeni su podaci o fizičko-hemijskim osobinama tla i specifičnim mikroklimatskim uslovima svake lokacije.

Eksperimenti su organizovani u okviru randomizovanog kompletnog blok dizajna sa tri ponavljanja. Osnovne parcele su bile podeljene na 4 reda, sa dužinom reda od 10 m. Razmak između redova bio je 70 cm, dok je razmak između biljaka u redu bio 25 cm. Podaci o prinosu semena su prikupljeni sa parcela iz srednjih redova kako bi se minimizirao rubni efekat. Merenje je izvršeno korišćenjem softvera *Easy Harvest* i *Harvest Master HM 800*, unutar kombajna *Wintersteiger Delta*. Sadržaj ulja u semenu je procenjivan pomoću NMR tehnike sa *Maran Ultra-10* analizatorom. Prinos semena je zabeležen u $t\ ha^{-1}$ na osnovu sadržaja vlage od 11%.

4.2.2 Opis skupa podataka

Podaci koji su korišćeni za predviđanje visine prinosa suncokretovog ulja bili su u obliku numeričkih vrednosti organizovani u strukturiranom tabelarnom formatu. Pošto su podaci prikupljeni tokom dve uzastopne godine, bilo je moguće pratiti performanse hibrida u promenljivim uslovima, uključujući različite vegetacione sezone, klimatske prilike i ekološke zone.

Podaci prikupljeni tokom obe godine uključivali su zapise za 1250 hibrida, pri čemu je za svaki hibrid sprovedeno po tri ponavljanja na sve četiri lokacije, što je rezultiralo ukupnim brojem od 15.000 zapisa po godini.

Prosečna sezonska klimatska varijabilnost na četiri testirana lokaliteta prikazana je u tabeli A.1. Prosečna temperatura je varirala od 13,2°C u aprilu do 24,5°C u avgustu, što ukazuje da je temperaturni režim bio u opsegu koji je povoljan za razvoj suncokreta (Vear, 2016). Vegetaciona sezona je u 2018. bila generalno toplija od 2019. godine, pri čemu su proleća bila toplija, dok su leta bila podjednako topla u obe godine. Količina padavina zavisila je kako od godine, tako i od lokaliteta. Na lokalitetu Vršac, tokom vegetacione sezone 2019. godine, zabeležen je najmanji nivo padavina, što potvrđuje izrazito sušne uslove u toj godini. Nasuprot tome, tokom 2018. godine, ista lokacija je zabeležila najvišu količinu kiše od svih lokaliteta, što je stvorilo znatno povoljnije uslove za rast biljaka. S druge strane, na lokalitetu Subotica situacija je bila potpuno drugačija. Godina 2019. obeležena je najvećom količinom padavina, dok je tokom 2018. godine na ovoj lokaciji palo najmanje kiše.

Kada je reč o uticaju klimatskih uslova na visinu prinosa ulja, 2018. godinu karakterisale su više temperature i manje padavine, ali je ukupan prinos ulja kod ispitivanih hibrida bio veći u 2019. Ovaj rezultat najpre deluje kontraintuitivno, međutim, time je naglašena važnost distribucije padavina, a ne samo ukupnih padavina. Raspodela padavina je u 2019. godini bila povoljnija za proizvodnju suncokreta, posebno u odnosu na ključne faze rasta i potrebe za unosom vode. Pošto biljke suncokreta imaju različite faze rasta sa različitim potrebama za vodom, vreme padavina igra ključnu ulogu u maksimizovanju prinosa ulja. Dobro raspoređeno snabdevanje vodom tokom ključnih faza, kao što su cvetanje i punjenje semena, može značajno povećati akumulaciju ulja, čak i u godinama sa nešto nižim ukupnim temperaturama.

Skup podataka je analiziran iz dva različita aspekta, na osnovu kojih su kreirana dva podskupa podataka. Prvi se fokusira na karakteristike hibrida suncokreta, uključujući njihove genetičke osobine i morfološke parametre. Drugi aspekt ispituje učinak ovih hibrida u različitim agroekološkim uslovima,

pružajući uvid u njihove performanse u zavisnosti od lokaliteta, odnosno od specifičnih klimatskih i zemljišnih faktora na datom lokalitetu.

U prvom podskupu podataka (tabela 8), najpre su izračunate srednje vrednosti prinosa semena za svaki hibrid u odnosu na tri replikacije, a zatim je izračunata prosečna vrednost prinosa svakog hibrida na svim lokacijama. Nakon toga, hibridi su kategorizovani u tri grupe u odnosu na vrednost prinosa semena: hibridi sa niskim, srednjim i visokim prinosom.

Na osnovu specifičnih osobina i namene, hibridi suncokreta su klasifikovani i u sledeće grupe:

- OIL – hibridi koji se prvenstveno koriste za proizvodnju ulja.
- IMI – hibridi otporni na herbicide iz grupe imidazolinona.
- SU – hibridi otporni na herbicide iz grupe sulfonilurea.
- HO – hibridi sa visokim sadržajem oleinske kiseline, idealni za specifične prehrambene industrije.
- CON – hibridi namenjeni za proizvodnju konditorskih proizvoda.

Za svaki hibrid je zabeležena rezistencija na volovod (*Orobanche cumana* - Or) i plamenjaču (*Plasmopara halstedii* - Pl), na osnovu prisustva specifičnih gena otpornosti. Takođe, zabeležena je klasa zrelosti hibrida, sa kategorijama u rasponu od ranog do kasnog sazrevanja.

Sadržaj ulja je izmeren korišćenjem tehnologije nuklearne magnetne rezonance (NMR), dok je prinos ulja izračunat na osnovu prinosa semena i procentualnog sadržaja ulja, prema standardnom pristupu korišćenom u radu Cvejić et al. (2023) primenom sledeće formule:

$$\text{Prinos ulja (kg/ha)} = \text{Prinos semena (kg/ha)} \times \frac{\text{Sadržaj ulja (\%)}}{100}$$

Tabela 8. Opis prvog podskupa podataka.

Karakteristika	Tip karakteristike	Broj grupa	Opis grupa
Prinos semena	Prediktor	3	Nizak, srednji, visok
Tip hibrida	Prediktor	5	OIL, IMI, SU, HO, CON
Otpornost na volovod	Prediktor	2	Bez Or gena, sa Or genima
Otpornost na plamenjaču	Prediktor	2	Bez Pl gena, sa Pl genima
Zrelost	Prediktor	3	Rani, srednji, kasni
Prinos ulja	Ciljno	-	-

U drugom podskupu podataka (tabela 9) fokus je na performansama hibrida na svakoj pojedinačnoj lokaciji, bez usrednjavanja rezultata. Ovaj podskup naglašava uticaj lokalnih faktora životne sredine (tipa zemljišta) i vremenskih uslova na visinu prinosa ulja. Prikupljeni podaci za obe godine su pažljivo analizirani kako bi se omogućila sveobuhvatna procena performansi hibrida u različitim ekološkim kontekstima.

Tabela 9. Opis drugog podskupa podataka.

Karakteristika	Tip karakteristike	Broj grupa	Opis grupa
Prinos semena	Prediktor	3	Nizak, srednji, visok
Lokalitet (tip zemljišta i meteorološki parametri)	Prediktor	4	SU, NS, KI, VŠ
Hibrid	Prediktor	1250	Tip hibrida (1-1250)
Prinos ulja	Ciljno	-	-

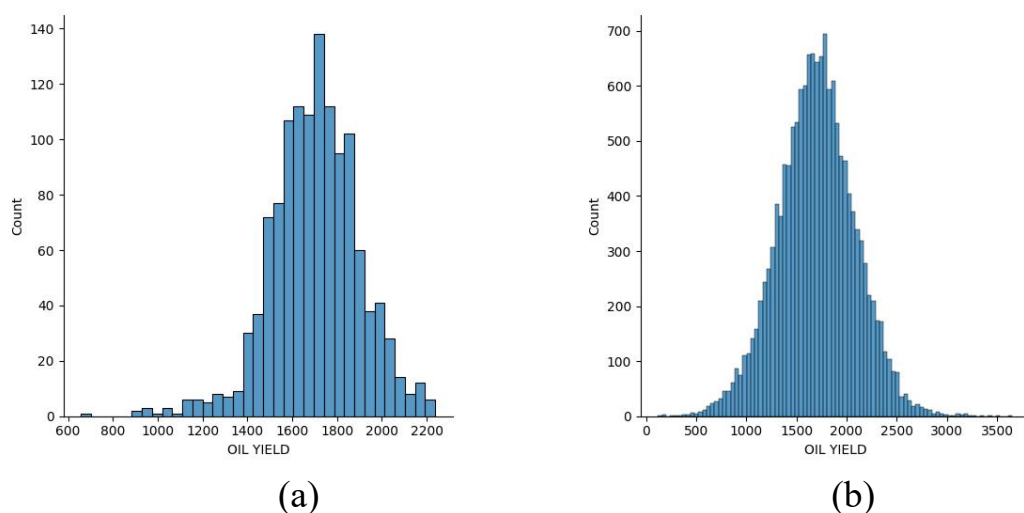
Nakon prikupljanja, svi podaci su prošli kroz proces prečišćavanja, sortiranja i preprocesiranja kako bi se osigurao njihov kvalitet i kako bi se pripremili za dalju upotrebu u modelima mašinskog učenja. Ovaj korak obuhvata uklanjanje nepotpunih, neispravnih ili redundantnih podataka, čime se povećava preciznost i pouzdanost budućih analiza. Takođe, izvršena je kategorizacija podataka, kojom su grupisane relevantne informacije, kao i standardizacija, koja je omogućila da podaci budu usklađeni u pogledu formata.

4.2.3 Eksploratorna analiza skupa podataka

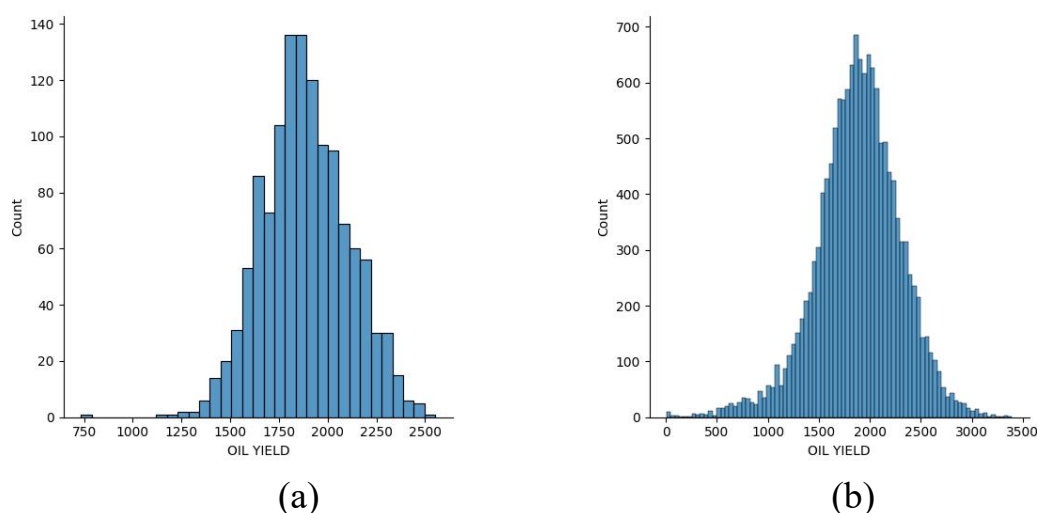
Prvi korak analize podataka u istraživanju bio je posmatranje skupova podataka u kontekstu nedostajućih vrednosti. Pažljivo rukovanje nedostajućim vrednostima važno je za održavanje kvaliteta i pouzdanosti skupa podataka, a njihovo zanemarivanje može dovesti do pogrešnih zaključaka. Skup podataka iz 2018. godine bio je potpun, dok je skup podataka iz 2019. godine sadržao nedostajuće vrednosti. Uzimajući u obzir ukupnu količinu podataka (koja je bila velika u odnosu na podatke koji su nedostajali), odlučeno je da se ti zapisi jednostavno uklone iz skupa podataka. Nakon uklanjanja, ukupan broj preostalih zapisa bio je 14699, odnosno uklonjen je 301 zapis sa nedostajućim vrednostima.

Nakon uklanjanja nedostajućih vrednosti iz skupa podataka, vizuelno je prikazana distribucija ciljne karakteristike. Slike 10 i 11 predstavljaju

vizuelizacije ciljne karakteristike za skup podataka iz 2018. i 2019. godine. Slike 10a i 11a ilustruju ukupnu prosečnu distribuciju prinosa ulja među hibridima, naglašavajući varijacije u performansama nezavisno od faktora specifičnih za lokaciju. Slike 10b i 11b, s druge strane, pokazuju distribuciju prinosa ulja za pojedinačne hibride na svakoj lokaciji, nudeći uvid u uticaj uslova vezanih za lokaciju kao što su tip zemljišta, padavine i temperatura.



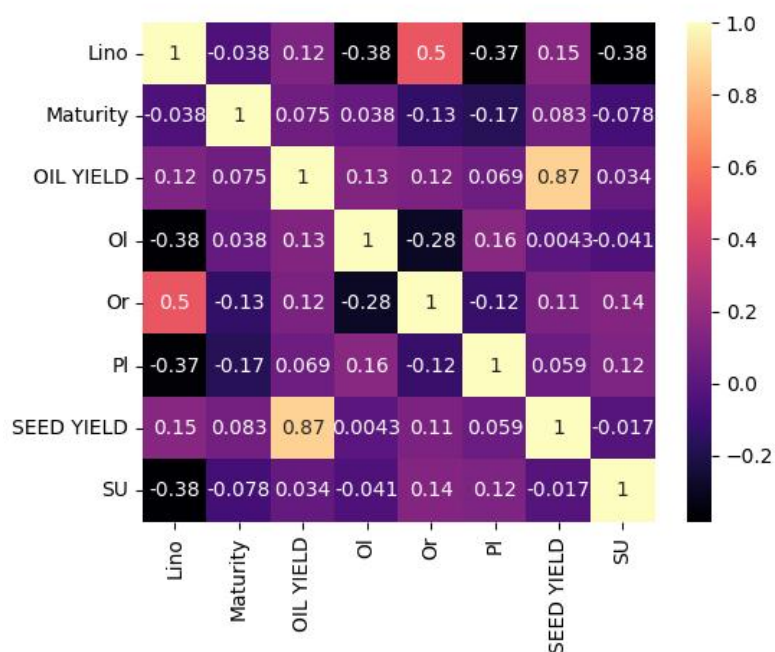
Slika 10. Distribucija prinosa ulja za skup podataka iz 2018. godine: (a) prvi podskup podataka; (b) drugi podskup podataka.



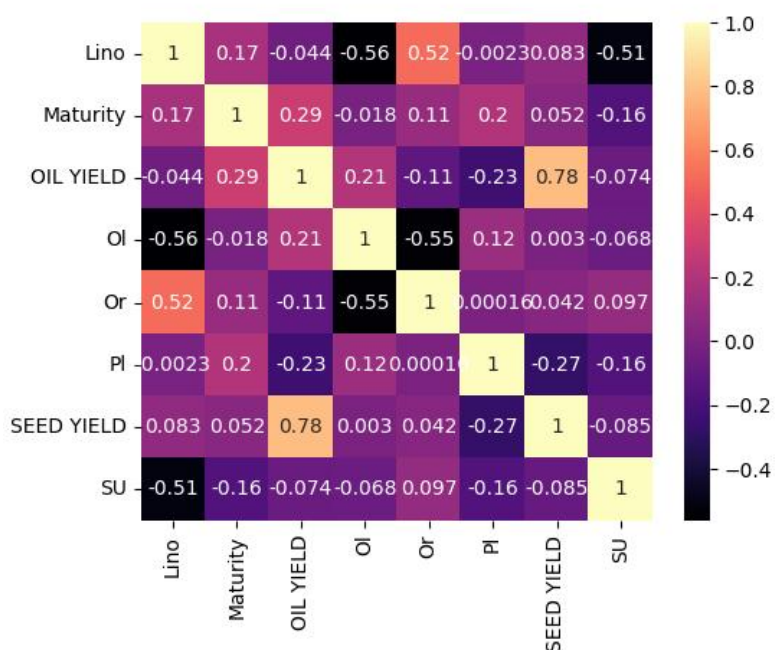
Slika 11. Distribucija prinosa ulja za skup podataka iz 2019. godine: (a) prvi podskup podataka; (b) drugi podskup podataka.

Pirsonovi koeficijenti korelacije (engl. *Pearson Correlation Coefficient*) predstavljaju metriku u statističkoj analizi koja kvantifikuje jačinu i smer linearne povezanosti između numeričkih varijabli. Pirsonovi koeficijenti korelacije između svih parova numeričkih varijabli prikazani su na slici 12 (za 2018. godinu) i slici 13 (za 2019. godinu). Ove korelacije su izračunate

isključivo za prvi podskup podataka jer on sadrži numeričke varijable, dok drugi podskup sadrži samo kategoričke varijable koje nisu pogodne za ovaj tip analize.



Slika 12. Koeficijenti Pirsonovih korelacija numeričkih varijabli za podatke iz 2018. godine.



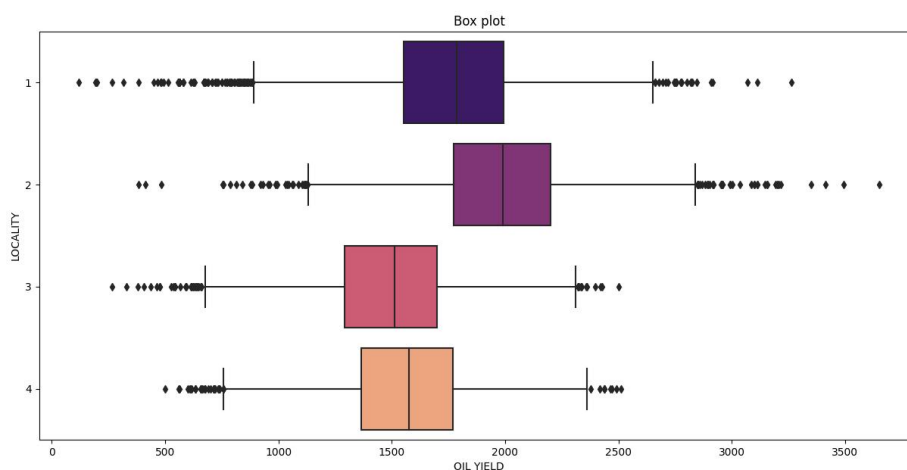
Slika 13. Koeficijenti Pirsonovih korelacija numeričkih varijabli za podatke iz 2019. godine.

Na osnovu slika 12 i 13, vidi se da postoji snažna pozitivna korelacija između prinosa ulja i prinosa semena, što je u skladu sa očekivanjima. Ovaj odnos je intuitivan jer prinos ulja direktno zavisi od prinosa semena i sadržaja ulja. Uočena korelacija naglašava prediktivnu snagu prinosa semena u proceni prinosa ulja.

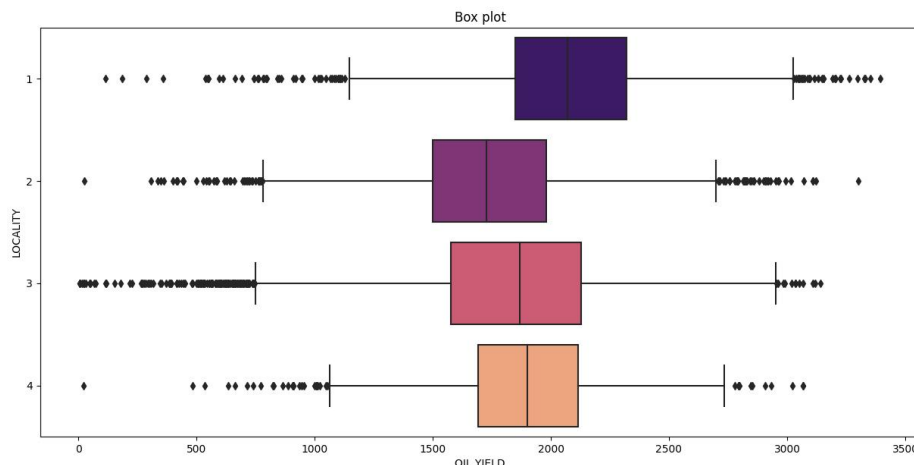
U istraživanju je ispitano da li karakteristike specifične za hibrid ili varijable specifične za lokaciju mogu poslužiti kao održivi prediktori prinosa ulja u scenarijima u kojima podaci o sadržaju ulja nisu dostupni. Sa agronomskog stanovišta, zanimljivo je razmotriti u kojoj meri navedene karakteristike mogu nadomestiti nedostatak direktnih merenja sadržaja ulja i da li u tom slučaju mogu precizno da predvide visinu prinosa ulja.

Korelacije između karakteristika su bile slične tokom obe godine, što ukazuje na stabilne odnose tokom vremena. Štaviše, analiza nije otkrila parove karakteristika sa preterano visokim korelacijama (izuzev prinosa semena i prinosa ulja), što bi ukazivalo na redundantnost varijabli. Kao rezultat toga, nijedna varijabla nije isključena iz početnog skupa podataka. Ovo obezbeđuje ravnotežu između maksimizovanja bogatstva skupa podataka i održavanja interpretabilnosti i efikasnosti prediktivnih modela.

Na slikama 14 i 15 nalaze se grafikoni koji prikazuju distribuciju vrednosti prinosa ulja za 2018. i 2019. godinu, u odnosu na četiri eksperimentalna lokaliteta. Sa grafikona na slikama se vidi da se prinos ulja razlikuje između lokaliteta unutar iste godine, kao i između dve godine na istom lokalitetu, što ukazuje na značajan uticaj lokalnih uslova i klimatskih faktora tokom vegetacionog perioda. Prosečne vrednosti prinosa ulja u 2019. godini bile su veće od onih zabeleženih 2018. godine. Ova razlika verovatno potiče iz varijacija u klimatskim uslovima između dve vegetacione sezone, kao što su raspored padavina i temperaturni režim tokom faza formiranja prinosa.



Slika 14. Vrednosti visine prinosa ulja u odnosu na lokalitet u 2018. godini.



Slika 15. Vrednosti visine prinosa ulja u odnosu na lokalitet u 2019. godini.

Detaljnije ispitivanje pokazalo je da je u 2018. godini najveći prinos ulja zabeležen na drugom lokalitetu (Novi Sad), dok je u 2019. godini prvi lokalitet (Subotica) imao najveći prinos. Za 2018. godinu vrednosti prinosa ulja za Suboticu bile su nešto ispod onih za Novi Sad, dok su vrednosti za Novo Miloševo i Vršac bile slične. Nasuprot tome, tokom 2019. godine vrednosti prinosa ulja postale su ujednačenije na svim lokalitetima, sa izuzetkom Subotice, u kojoj je postignuta najveća stabilnost u prinosima.

Iako su eksperimentalne lokacije geografski blizu, one pokazuju značajne varijacije u tipu zemljišta i mikroklimatskim uslovima, koji značajno utiču na performanse useva. Subotica ima dosledne rezultate u prinosu ulja u obe godine, što znači da su hibridi testirani na ovoj lokaciji dobro prilagođeni njenom peskovitom zemljištu. Peskovito zemljište, koje karakterišu brza drenaža vode i manji kapacitet zadržavanja vode, nameću stres biljkama, primoravajući biljke suncokreta da razviju dublje korenove sisteme kako bi pristupili vlazi. Čini se da je ova adaptacija ublažila efekte fluktuacije padavina i temperature, obezbeđujući stabilne prinose u Subotici.

Ova zapažanja naglašavaju važnost razmatranja tipa zemljišta i mikroklimatskih faktora pri odabiru i kultivaciji hibrida suncokreta. Rezultati iz Subotice ističu potencijal peskovitog zemljišta za uzgoj suncokreta, posebno kada se koriste hibridi sposobni za dubok razvoj korena, što ih čini manje osetljivim na prolazne deficite vode.

4.2.4 Korišćeni modeli mašinskog učenja

U istraživanju su korišćeni različiti modeli mašinskog učenja kako bi se njihove performanse uporedile i obezbedila sveobuhvatna i efikasna analiza

rezultata. Izbor modela zasnovan je na njihovoj sposobnosti da obrade različite tipove podataka i prepoznaju složene obrasce, kao i činjenici da su se pokazali uspešnim u sličnim oblastima primene.

Korišćeni su metod potpornih vektora za regresiju, metod K-najbližih suseda, veštačka neuronska mreža i metod slučajne šume.

4.2.5 Obuka modela mašinskog učenja

Inicijalno obučavanje modela mašinskog učenja sprovedeno je da bi se utvrdila optimalna konfiguracija hiperparametara koja omogućava najvišu tačnost predviđanja visine prinosa ulja. Optimizacija hiperparametara sprovedena je radi poboljšanja generalizacione sposobnosti modela i smanjenja rizika od prekomernog prilagođavanja.

Određivanje optimalne vrednosti k za K-NN algoritam sprovedeno je nezavisno za svaki podskup podataka identifikacijom najviše vrednosti koeficijenta determinacije R^2 u rasponu od k vrednosti. Konkretno, testirane vrednosti k kretale su se od 2 do 30 kako bi se obezbedila što sveobuhvatnija pretraga najpogodnijeg parametra. Za prvi podskup podataka, koji pokriva obe godine analize, optimalna vrednost k je određena na osnovu minimalne greške predikcije i iznosila je 12. Za drugi podskup podataka, optimalna vrednost je bila značajno veća i iznosila je 25. Ovim se vidi varijabilnost vrednosti k u zavisnosti od karakteristika skupa podataka.

Za model veštačke neuronske mreže najbolji hiperparametri su identifikovani korišćenjem tehnike unakrsne validacije pretrage mreže, koja je implementirana u biblioteci *scikit-learn*. Ovaj metod sistematski procenjuje kombinacije hiperparametara da bi se odredila konfiguracija koja daje najbolje performanse. Optimizator koji je korišćen tokom obuke za ANN bio je Adam, popularan izbor zbog svoje efikasnosti i mogućnosti prilagodljive brzine učenja.

U pogledu konfiguracije obuke, optimalni parametri su varirali između podskupova podataka i godina. Za prvi podskup podataka iz 2018. godine, optimalan broj epoha obuke je bio 20, sa veličinom serije od 20. Za odgovarajući prvi podskup podataka iz 2019. model je bio najbolji kada je obučen sa 30 epoha i veličinom serije od 10.

Isti proces primenjen je i za drugi podskup podataka. Za podatke iz 2018. godine, optimalan broj epoha je bio 30, sa veličinom serije od 10. Nasuprot tome, za podatke iz 2019. godine, obuka je zahtevala 50 epoha sa veličinom serije od 20 da bi se postigli optimalni rezultati.

Pored toga, arhitektura ANN-a je uključivala skriveni sloj sa 5 neurona. Ova vrednost bila je konzistentna u svim podskupovima podataka za sve godine.

Korišćeni parametri obezbedili su da je obuka neuronske mreže prilagođena specifičnim karakteristikama svakog skupa podataka, smanjujući rizik od prekomernog prilagođavanja i omogućavajući pouzdanu procenu prinosa ulja na osnovu karakteristika hibrida i lokalnih faktora.

4.2.6 Metode evaluacije

Za evaluaciju modela korišćeno je više metrika koje omogućavaju sveobuhvatnu analizu performansi modela, pružajući uvid u preciznost predikcija, sposobnost modela da generalizuje, kao i njegovu robusnost.

Korišćene su: srednja apsolutna greška (engl. *Mean Absolute Error*, MAE), srednja kvadratna greška (engl. *Mean Squared Error*, MSE), koren srednje kvadratne greške (engl. *Root Mean Squared Error*, RMSE) i koeficijent determinacije (engl. *R-Squared*, R^2).

MAE je suma apsolutnih razlika između predviđenih i stvarnih vrednosti podeljena brojem posmatranja (jednačina 5).

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (\text{Jednačina 5})$$

MSE je suma kvadrata razlika između predviđenih i stvarnih vrednosti podeljena brojem posmatranja (jednačina 6).

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (\text{Jednačina 6})$$

RMSE je kvadratni koren MSE, što ukazuje na standardnu devijaciju reziduala odnosno greške predviđanja (jednačina 7).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (\text{Jednačina 7})$$

R^2 je proporcija varijanse u zavisnoj promenljivoj koju model može da objasni na osnovu prediktora (jednačina 8).

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (\text{Jednačina 8})$$

U prethodnim jednačinama su upotrebljene sledeće oznake:

- n - broj uzoraka,
- y - stvarna vrednost prinosa ulja,
- $\hat{y}$ - predviđena vrednost prinosa ulja,
- $\bar{y}$ - srednja vrednost prinosa ulja.

4.3 Implementacija softverskog rešenja

Sve analize opisane u radu sprovedene su korišćenjem programskog jezika *Python*, verzija 3.6 (Python Software Foundation, Wilmington, DE, USA), koji je poznat po velikoj fleksibilnosti i širokoj podršci za statističke analize i mašinsko učenje. Takođe, *Python* pruža bogat ekosistem biblioteka koje omogućavaju efikasnu implementaciju, analizu i vizualizaciju podataka, što ga čini dobrim izborom za ovu studiju.

Za implementaciju neuronskih mreža korišćen je *TensorFlow* okvir (verzija 2.4.1) kao *backend* i *Keras* (verzija 2.4.3) kao *frontend*. *TensorFlow* je odabran zbog svoje brzine, skalabilnosti i sposobnosti da podrži složene arhitekture dubokog učenja. *Keras*, kao API visokog nivoa, olakšava definisanje i obuku neuronskih mreža zahvaljujući svojoj intuitivnoj sintaksi i integraciji sa *TensorFlow*-om.

Ključne *Python* biblioteke koje su korišćene u implementacionom delu istraživanja su: *Pandas*, *NumPy*, *scikit-learn*, *Matplotlib* i *Seaborn*.

Biblioteke *Pandas* i *NumPy* su korišćene u fazi pripreme i obrade podataka. *Pandas* je biblioteka koja se koristi za manipulaciju i analizu podataka, uključujući rad sa strukturama podataka poput *DataFrame* objekata, koji omogućavaju jednostavno upravljanje velikim skupovima podataka.

NumPy se koristi za numeričke operacije, posebno one koje uključuju rad sa višedimenzionalnim nizovima i matematičkim funkcijama visokih performansi.

Biblioteka *scikit-learn* je korišćena za implementaciju algoritama mašinskog učenja, podelu na trening i test skupove, standardizaciju podataka, optimizaciju hiperparametara, kao i metrike evaluacije.

Biblioteke *Matplotlib* i *Seaborn* su upotrebljene za vizualizaciju podataka. *Matplotlib* je korišćen za kreiranje standardnih grafikona, dok je *Seaborn*, koji sadrži ugrađene stilove i kompleksne funkcije, upotrebljen za napredne statističke vizuelizacije.

Radi obezbeđivanja dodatnih funkcionalnosti, korišćene su i biblioteke *Statsmodels* i *OpenCV*. Biblioteka *Statsmodels* je omogućila primenu naprednih statističkih analiza, regresionih modela i testiranje hipoteza. Za obradu slika je korišćena biblioteka *OpenCV*, pre svega za pripremu i augmentaciju ulaznih podataka u vidu slika.

5. Rezultati

U okviru ovog poglavlja predstavljeni su rezultati eksperimenata iz prethodnog poglavlja.

5.1. Rezultati detekcije stresa usled suše

5.1.1 Rezultati eksperimenta u inicijalnoj postavci

U inicijalnoj postavci eksperimenta upoređene su performanse za pet odabranih konvolucionih neuronskih mreža radi procene njihove efikasnosti u detekciji stresa usled suše na slikama biljaka. Ispitane su različite arhitekture modela, njihova tačnost i stabilnost prilikom obuke, kao i uticaj formata slike (u boji i sivi tonovi) na konačne rezultate. Analizirana je vremenska efikasnost svake arhitekture, kao i ponašanje modela tokom obuke. Na osnovu ovih rezultata identifikovan je model koji pokazuje najvišu stabilnost, preciznost i potencijal za dalja eksperimentisanja.

5.1.1.1 Poređenje performansi modela i uticaja formata slika

Prva grupa eksperimenata obuhvatala je korišćenje pet različitih arhitektura konvolucionih neuronskih mreža, na originalnim slikama i na njihovim sivim verzijama (u *grayscale* formatu). Rezultati dobijeni korišćenjem ovih pet arhitektura su upoređeni i predstavljeni u tabeli 10. U tabeli su za svaku kombinaciju CNN arhitekture i vrstu slika, prikazani tačnost, preciznost, odziv, F1-skor, test gubitak i prosečno vreme izvršavanja po epohi.

Tabela 10. Poređenje performansi pet odabranih CNN modela (najbolje vrednosti su podebljane).

CNN Model	Boja	Tačnost	Preciznost	Odziv	F1-skor	Test Gubitak	Vreme/Ep. (s)
VGG16	U boji	0,82	0,81	0,86	0,84	0,82	250,81
	Siva	0,65	0,62	0,84	0,72	0,65	233,69
VGG19	U boji	0,66	0,91	0,39	0,55	0,66	306,31
	Siva	0,67	0,81	0,49	0,61	0,67	282,62
DenseNet 121	U boji	0,85	0,78	1,00	0,87	0,85	103,00
	Siva	0,93	0,98	0,88	0,93	0,93	90,91
Inception V3	U boji	0,85	0,78	1,00	0,87	1,61	69,16
	Siva	0,93	0,94	0,92	0,93	0,13	51,28

MobileNet	U boji	0,91	0,92	0,90	0,91	0,19	56,60
	Siva	0,89	0,88	0,90	0,89	0,22	33,29

Najbolji rezultati za skup podataka sa slikama u boji postignuti su kada je korišćena MobileNet arhitektura. Tada su tačnost i F1-skor bili 0,91. Ovaj model je takođe bio efikasan i u slučaju kada su za skup podataka korišćeni sivi tonovi slika i tada su tačnost i F1-skor bili 0,89.

Sa DenseNet arhitekturom u kombinaciji sa sivim slikama, postignuti su tačnost i F1-skor od 0,93 što je prevazišlo performanse za originalni format slika, kada je tačnost bila 0,85, a F1-skor 0,87. Slični rezultati dobijeni su i sa Inception modelom. Tada su postignuti visoka vrednost tačnosti i F1-skor od 0,93 za skup podataka sa slikama u sivom formatu. Međutim, korišćenjem InceptionV3 modela na skupu podataka sa slikama u boji, vrednosti su bile niže - 0,85 za tačnost i 0,87 za F1-skor.

Oba VGG modela su postigla značajno slabije rezultate kada je reč o tačnosti i F1-skor u odnosu na sve ostale arhitekture. VGG16 je postigao nešto bolje rezultate za skup podataka sa slikama u boji (tačnost 0,82, F1-skor 0,84), dok su za skup podataka sa slikama u sivom formatu rezultati bili znatno lošiji (tačnost 0,65, F1-skor 0,72). Takođe, VGG19 je imao čak i nižu tačnost za oba formata (0,66 za slike u boji i 0,67 za sive slike), uz znatno manji F1-skor od ostalih modela.

Kada je reč o ostalim metrikama, DenseNet model za slike u boji ima veoma visok odziv (1,00), ali nešto nižu preciznost (0,78). U sivom formatu, DenseNet ima najbolju preciznost od svih modela (0,98) i veoma dobar odziv (0,88). Slične rezultate postigao je i InceptionV3 model, sa odzivom od 1,00 i preciznošću od 0,78 za format u boji i 0,92 odzivom i 0,94 preciznošću za sive slike.

VGG modeli su imali nesklad između preciznosti i odziva. VGG19 za originalni skup podataka ima visoku preciznost (0,91), ali vrlo nizak odziv (0,39), što ukazuje na problem sa balansiranjem predikcija. S druge strane, kod slika u sivim tonovima, preciznost i odziv su nešto izjednačeniji, ali su i dalje značajno niži od drugih arhitektura. To sugeriše da VGG modeli nisu optimalni za ovaj zadatak i da druge CNN arhitekture, poput DenseNet-a i InceptionV3, značajno bolje generalizuju podatke i pružaju konzistentnije rezultate.

MobileNet arhitektura u kombinaciji sa originalnim formatom slika, imala je visoku preciznost (0,92) i odziv (0,90), a ove vrednosti ostale su stabilne i u sivim tonovima (preciznost od 0,88 i odziv od 0,90).

5.1.1.2 Vremenska efikasnost modela

Prosečno vreme obučavanja modela po epohi je značajno variralo u odnosu na korišćeni CNN model, pri čemu je, očekivano, uočeno duže vreme obučavanja za slike u boji u odnosu na slike u sivom formatu.

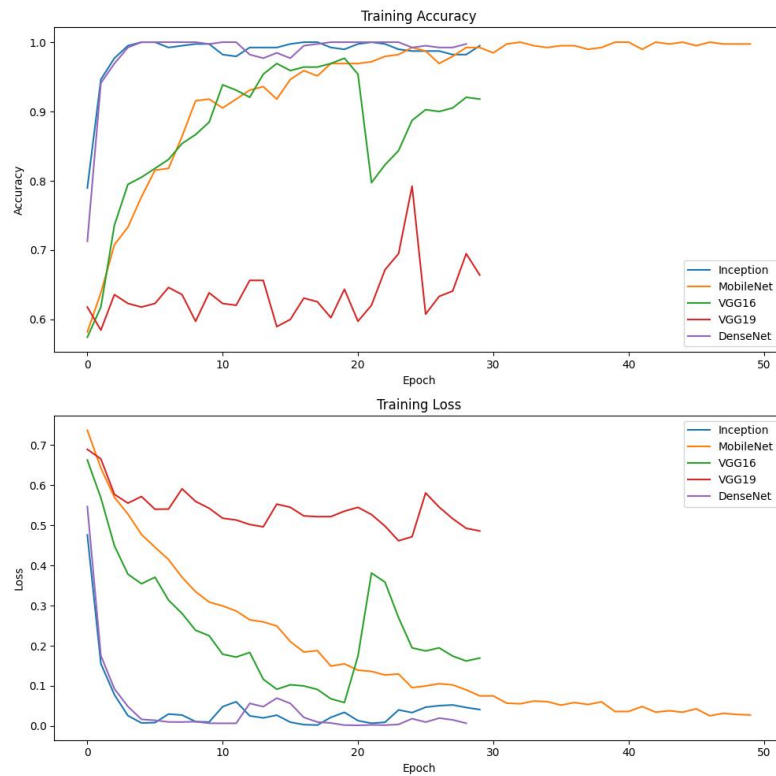
MobileNet se pokazao kao vremenski najefikasniji model, sa prosečnim vremenom obuke od 56,6 sekundi po epohi za slike u boji i 33,29 sekundi po epohi za sive tonove slika. InceptionV3 je imao slične performanse, sa 69,19 sekundi po epohi za skup podataka u boji i 51,28 sekundi za slike u sivim tonovima. Sledi model DenseNet, koji je zahtevao 103 sekunde po epohi za obuku nad skupom u boji i 90,91 sekundi po epohi za slike u sivim tonovima.

VGG modeli su bili najmanje efikasni, sa najdužim vremenom obuke u odnosu na druge arhitekture. Obuka VGG16 modela trajala je 250,81 sekundi po epohi za skup podataka u boji, a za sive tonove, trajala je 233,69 sekundi. Na kraju, VGG19 kao najmanje efikasan, zahtevao je 306,31 sekundi po epohi za skup podataka u boji i 282,62 sekundi po epohi za slike u sivim tonovima.

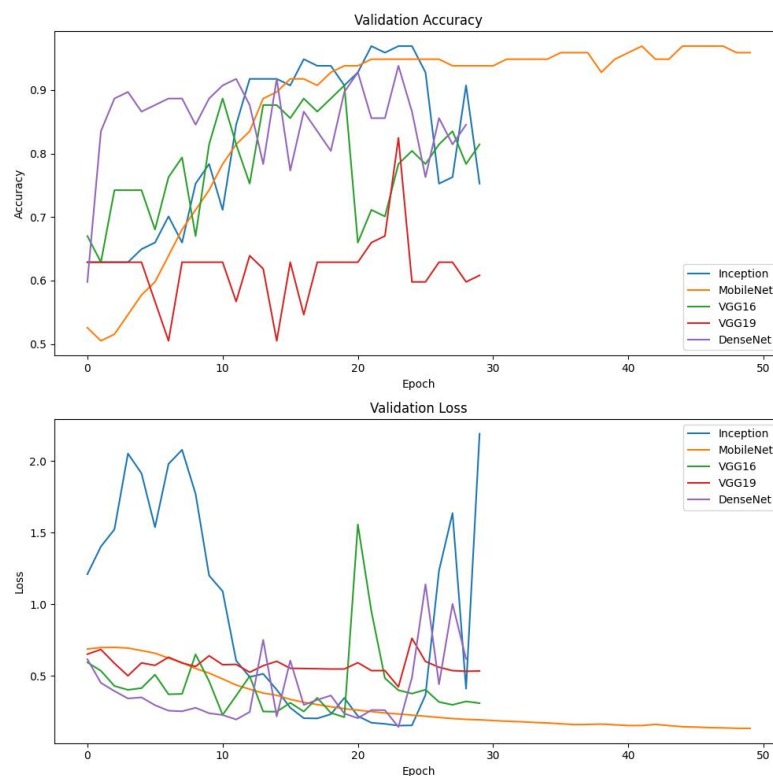
5.1.1.3 Vizuelna analiza obuke modela

Performanse modela su detaljno analizirane i uz pomoć vizuelnog prikaza krivih tačnosti i gubitka tokom obučavanja, za sve modele, nad trening i validacionim skupovima podataka.

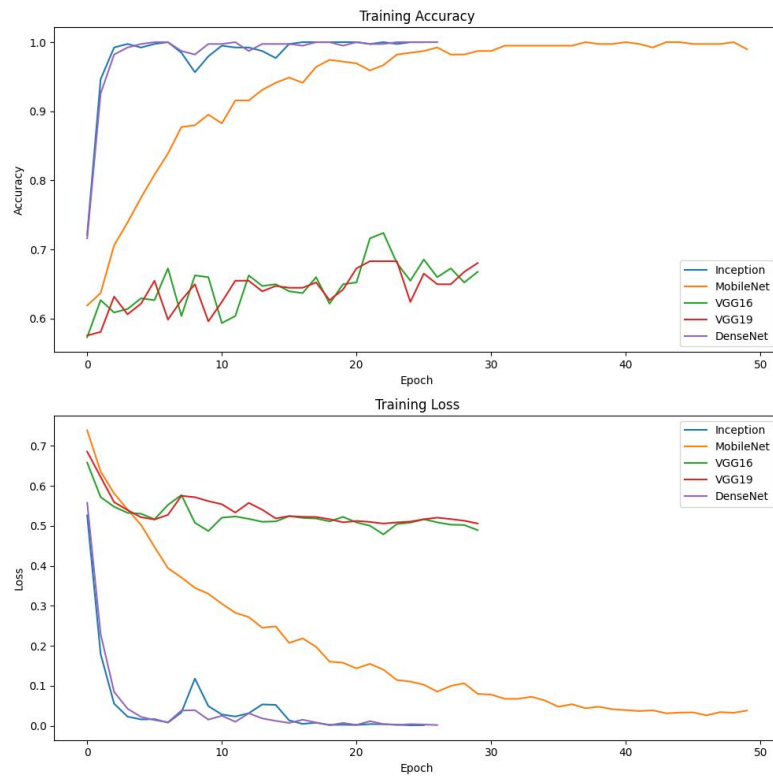
Slike 16 i 17 prikazuju kako su se tačnost i gubitak menjali tokom epoha za svaki model, za slike u boji. Na slikama 18 i 19 vide se promene tačnosti i gubitka za slike u sivom formatu.



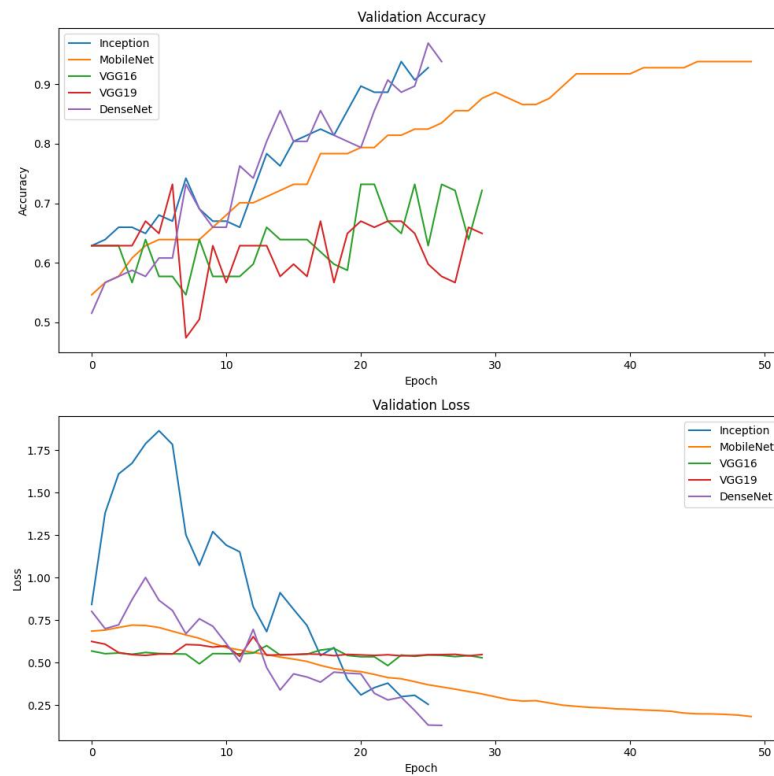
Slika 16. Rezultati tačnosti i gubitka za deo skupa podataka za obuku u boji.



Slika 17. Rezultati tačnosti i gubitka za deo skupa podataka za validaciju u boji.



Slika 18. Rezultati tačnosti i gubitka za deo skupa podataka za obuku u sivim tonovima.



Slika 19. Rezultati tačnosti i gubitka za deo skupa podataka za validaciju u sivim tonovima.

5.1.1.4 Analiza performansi i izbor optimalnog modela

VGG modeli VGG16 i VGG19, pokazali su najslabije performanse i nisku stabilnost na obe varijante skupa podataka. InceptionV3 je pokazao nedosledne rezultate, sa приметnim oscilacijama u tačnosti validacije i visokim gubitkom na originalnom skupu podataka. DenseNet je ostvario visoku tačnost uz nizak i stabilan gubitak tokom obuke.

Međutim, MobileNet je bio model koji je imao najkraće vreme obučavanja i najstabilnije vrednosti gubitka i tačnosti na validacionom skupu kroz epohe, zbog čega je odabran za nastavak istraživanja i dalja eksperimentisanja.

5.1.2 Evaluacija performansi MobileNet modela uz različite pristupe augmentacije podataka

Nakon identifikacije MobileNet modela kao optimalnog, sprovedena je evaluacija njegovih performansi prilikom obučavanja na slikama u boji i sivim tonovima, uz primenu različitih pristupa augmentacije podataka. Korišćena su tri različita pristupa:

1. bez augmentacije,
2. sa standardnim augmentacionim tehnikama,
3. sa ciljanom augmentacijom.

Numerički rezultati dobijeni primenom predloženih konfiguracija modela su prikazani u tabeli 11.

Tabela 11. Poređenje performansi MobileNet modela treniranog na tri varijante skupa podataka (najbolje vrednosti su podebljane).

Boja	Augmentacija podataka	Tačnost	Preciznost	Odziv	F1-skor	Test Gubitak
U boji	Bez augmentacije	0,91	0,92	0,90	0,91	0,19
	Standardna augmentacija	0,73	0,88	0,57	0,69	0,58
	Ciljana augmentacija	0,92	0,98	0,86	0,92	0,21
Sivi tonovi	Bez augmentacije	0,89	0,88	0,90	0,89	0,22
	Standardna augmentacija	0,74	0,86	0,61	0,71	0,47
	Ciljana augmentacija	0,95	1,00	0,90	0,95	0,17

5.1.2.1 Rezultati za slike u boji

Kada je model treniran na originalnom skupu podataka bez augmentacije, postignuti su veoma dobri rezultati, uključujući visoku tačnost (0,91), preciznost (0,92), osetljivost (0,90) i F1-skor (0,91). To ukazuje na sposobnost modela da dobro generalizuje na osnovu osnovnog skupa podataka.

Međutim, kada su primenjene standardne tehnike augmentacije podataka, došlo je do značajnog smanjenja tačnosti i odziva – tačnost je pala na 0,73, dok je odziv smanjen na 0,57. Ipak, preciznost je ostala visoka (0,88), što znači da je model i dalje bio sposoban da generiše tačne predikcije, ali sa većim brojem promašaja u detekciji pozitivne klase. Pretpostavlja se da su slabije performanse posledica prevelikih distorzija uvedenih standardnom augmentacijom.

S druge strane, primena ciljane augmentacije na originalne slike dovela je do poboljšanja ukupnih performansi modela. Tačnost i F1-skor su porasli na 0,92, što dokazuje da MobileNet uspešno uči iz varijacija koje unosi ciljane augmentacija.

Ovi rezultati ukazuju na to da dobro osmišljena augmentacija može doprineti boljoj generalizaciji modela.

5.1.2.2 Rezultati za slike u sivim tonovima

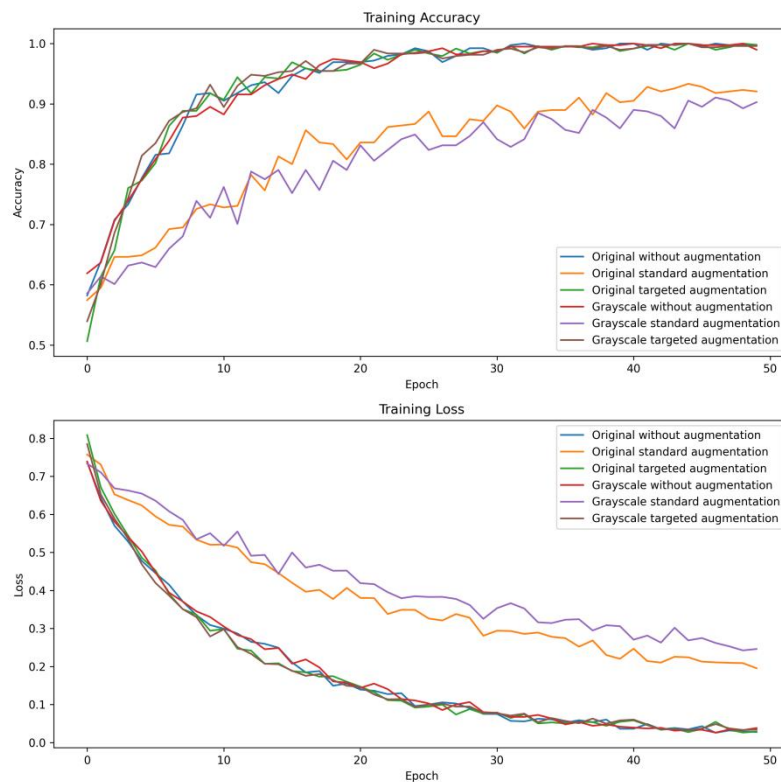
Kod sivih slika, MobileNet je postigao nešto slabije rezultate kada se koristio skup podataka bez augmentacije. Tačnost i F1-skor su iznosili 0,89, što ukazuje na to da model nije bio toliko efikasan u prepoznavanju relevantnih obrazaca na ovom tipu slika. Kada je korišćen skup podataka sa standardnom augmentacijom, tačnost je pala na 0,74, a osetljivost na 0,61, što ukazuje na smanjenu sposobnost modela da detektuje relevantne slučajeve.

Međutim, primena ciljane augmentacije opisane u metodologiji dovela je do značajnog poboljšanja performansi modela, koji je nadmašio sve druge konfiguracije modela sa tačnošću od 0,95, preciznošću od 1,0 i F1-skorom od 0,95. Ovaj rezultat ne samo da je nadmašio prethodne pokušaje, već je generisao i najniži test gubitak (0,17), što pokazuje visoku efikasnost modela pri primeni ciljane augmentacije.

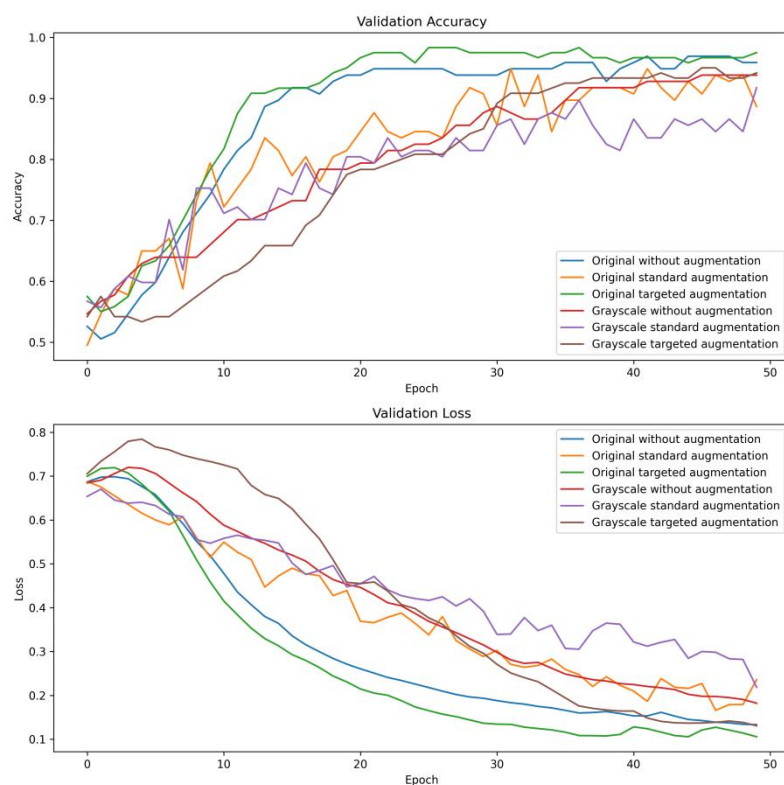
5.1.3 Vizuelna analiza rezultata MobileNet modela

Dobijeni rezultati su dodatno potvrđeni analizom funkcije gubitka koja je prikazana na slici 20. Na njoj se vidi da model postiže izuzetne performanse kada se koristi ciljana augmentacija, što potvrđuje sposobnost modela da u toj konfiguraciji podataka efikasno uči.

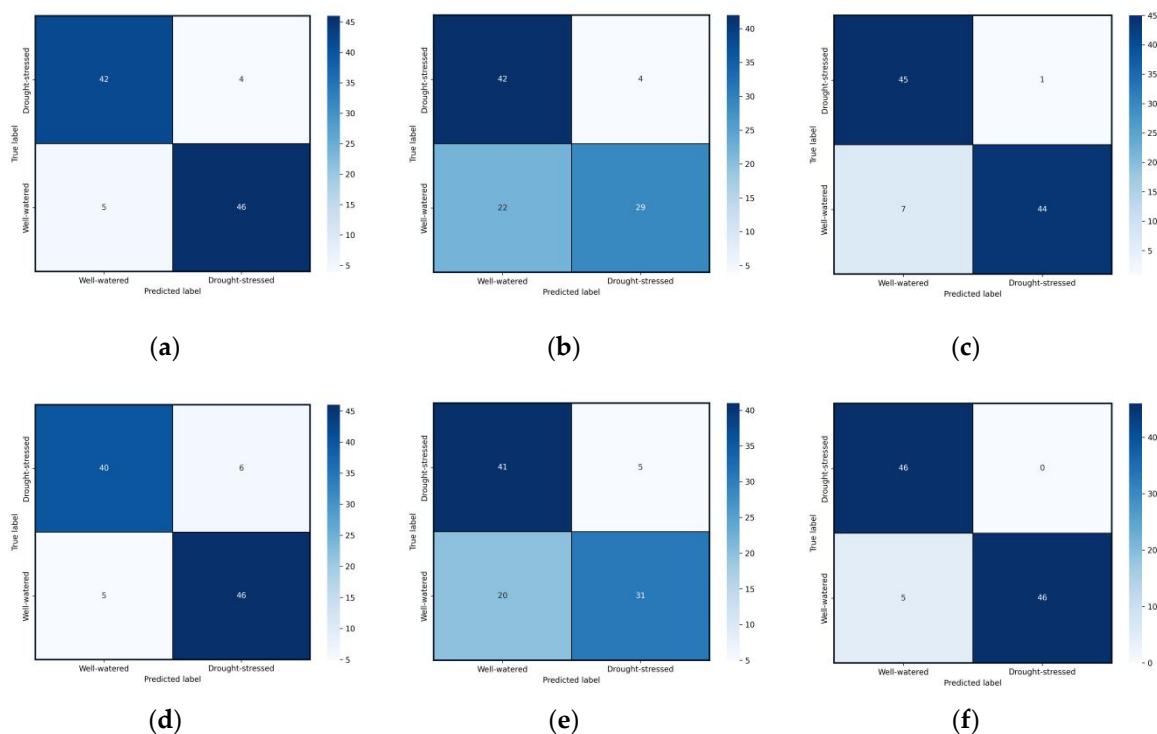
Na slikama 20 i 21 vidi se kako ciljana augmentacija podataka deluje na rezultate, pre svega na tačnost i gubitak, kod slika u boji i slika u sivim tonovima. Dodatna raznovrsnost i složenost koje uvodi ciljana augmentacija podataka mogu biti uzrok boljih performansi. Nedostatak dubine i varijacija boja u sivim slikama verovatno doprinosi boljim rezultatima, jer omogućava modelu da se više fokusira na strukturalne karakteristike, bez nepotrebnih varijacija povezanih sa bojama.



Slika 20. Krive tačnosti i funkcije gubitka MobileNet modela tokom obuke za originalni, augmentovani i ciljano augmentovani skup podataka.



Slika 21. Krive tačnosti i funkcije gubitka MobileNet modela tokom validacije za originalni, augmentovani i ciljano augmentovani skup podataka.



Slika 22. Matrice konfuzije koje prikazuju performanse različitih MobileNet konfiguracija na test skupu podataka.

Slika 22a prikazuje rezultate modela treniranog na originalnim podacima, odnosno podacima bez augmentacije. U ovom slučaju, model pokazuje zadovoljavajuću tačnost sa samo nekoliko pogrešnih klasifikacija, što ukazuje na to da osnovni skup podataka već sadrži dovoljno informacija za prepoznavanje razlika između klasa. Ipak, postoji blaga neuravnoteženost u tačnosti klasifikacije između klasa.

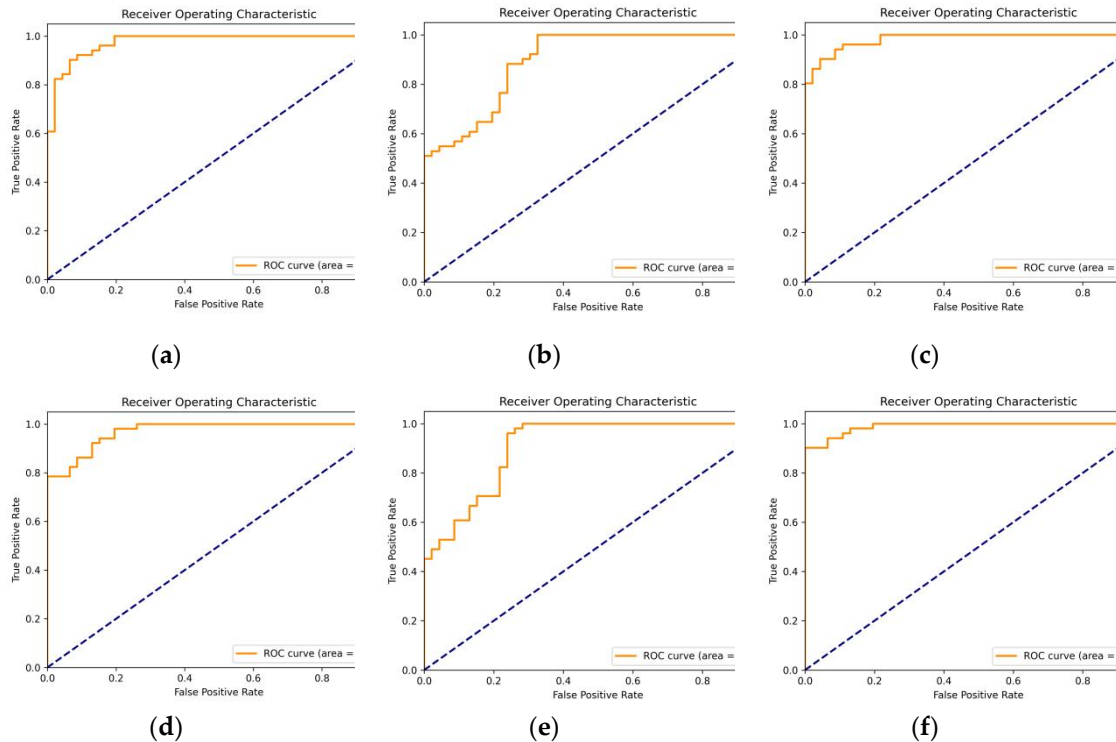
Druga matrica, prikazana na slici 22b, predstavlja rezultate modela koji koristi skup podataka u boji na kojem je primenjena standardna augmentacija podataka. Iako tačnost klasifikacije biljaka pod stresom ostaje na sličnom nivou, značajno se povećava broj pogrešno klasifikovanih dobro zalivenih biljaka. To znači da standardna augmentacija može uvesti dodatne varijacije koje model ne interpretira ispravno što smanjuje preciznost modela u razlikovanju klasa.

Slika 22c prikazuje rezultate modela treniranog sa slikama u boji i ciljanom augmentacijom. U ovom slučaju, broj pogrešnih klasifikacija je najmanji, što ukazuje na to da je ovaj pristup najefikasniji u očuvanju ključnih osobina klasa tokom augmentacije. Ovi rezultati potvrđuju da dobro osmišljena i specifično prilagođena augmentacija može značajno poboljšati tačnost modela.

Na slici 22d prikazana je matrica konfuzije za model treniran na slikama u sivim tonovima, bez augmentacije. Dobijeni rezultati su slični onima prikazanim na slici 22a. Model pokazuje solidnu tačnost, ali postoji blago neuravnotežen odnos između tačne klasifikacije. Model blago favorizuje klasifikaciju dobro zalivenih biljaka.

Uvođenje standardne augmentacije u slike u sivom formatu dovodi do povećane konfuzije između klasa, što se može videti na slici 22e. Slično kao i na slici 22b, značajno se povećava broj pogrešno klasifikovanih dobro zalivenih biljaka. Ovi rezultati ukazuju na smanjenje ukupne tačnosti modela i lošiju sposobnost razlikovanja biljaka u različitim uslovima.

Konfiguracija u kojoj se koriste ciljano augmentovane slike u sivom formatu, daje najbolje rezultate, što se vidi na matrici konfuzije na slici 22f. Ova konfiguracija ostvaruje najbolje performanse, sa 49 tačno klasifikovanih dobro zalivenih biljaka i 46 tačno klasifikovanih biljaka koje su bile izložene sušnom stresu. Samo 5 dobro zalivenih biljaka je pogrešno klasifikovano, dok su sve sušne biljke ispravno prepoznate. Ovo sugerise da ciljana augmentacija kod slika u sivom formatu značajno poboljšava razlikovanje biljaka izloženih različitim uslovima u poređenju sa standardnom augmentacijom i originalnim podacima.



Slika 23. Kriva operativnih karakteristika prijemnika za različite konfiguracije MobileNet modela.

Rezultati prikazani na krivama karakteristika operativnosti prijemnika (engl. *Receiver operating characteristic*, ROC) za različite konfiguracije MobileNet modela ukazuju na varijacije u performansama u zavisnosti od vrste podataka i primenjenih tehnika augmentacije. Model treniran na originalnim podacima (slika 23a) pokazuje zadovoljavajuću efikasnost modela, dok standardna augmentacija (slika 23b) ne dovodi do značajnog poboljšanja performansi. Najbolji rezultat postiže model sa ciljanom augmentacijom na podacima u boji (slika 23c), gde je AUC-vrednost (engl. *Area Under the Curve*) najviša ($\sim 0,98$), što sugerise da ciljana augmentacija podataka značajno doprinosi preciznosti modela.

Kada su podaci konvertovani u sive tonove (slike 23d, 23e i 23f), primetno je blago smanjenje performansi. Model treniran na neaugmentovanim podacima u sivim tonovima (slika 23d) pokazuje nižu AUC vrednost u poređenju sa originalnim podacima, što ukazuje na značaj informacija o bojama u procesu klasifikacije. Standardna augmentacija podataka u sivim tonovima (slika 23e) ne donosi značajna poboljšanja, dok ciljana augmentacija (slika 23f) poboljšava sposobnost klasifikacije, ali ne dostiže performanse ostvarene na podacima u boji.

Dobijeni rezultati ukazuju na to da ciljana augmentacija značajno doprinosi poboljšanju modela, posebno kada se koristi na slikama u boji. S druge strane,

konverzija slika u sive tonove smanjuje performanse, verovatno zbog gubitka ključnih vizuelnih informacija koje boja pruža.

5.2. Rezultati predikcije visine prinosa ulja

5.2.1. Performanse regresionih algoritama

U ovom poglavlju predstavljeni su rezultati evaluacije modela za predikciju visine prinosa suncokretovog ulja dobijeni primenom četiri regresiona algoritma navedena u metodologiji. Analiza performansi modela sprovedena je na podacima iz dve eksperimentalne sezone (2018. i 2019. godina) koji su posmatrani kao podskupovi podataka. Na oba podskupa podataka poređene su tačnost predikcije i stabilnost modela.

Za evaluaciju performansi upotrebljene su standardne regresione metrike (MAE, MSE, RMSE i R^2). Rezultati pokazuju da su performanse modela značajno varirale u zavisnosti od izabranog algoritma i karakteristika posmatrane godine.

Model zasnovan na metodi slučajne šume (RFR) nadmašio je druge modele u gotovo svim metrikama u obe godine, što ukazuje na to da predstavlja najefikasniji pristup za predviđanje visine prinosa suncokretovog ulja.

Korišćenjem metode k-najbližih suseda (K-NN) dobijeni su rezultati odmah iza RFR modela. Da bi se odredio optimalan broj suseda (k) za K-NN, sprovedena je evaluacija za vrednosti parametra k u opsegu od 2 do 30 posebno za svaki podskup podataka u obe godine. Kriterijum izbora zasnivao se na maksimizovanju rezultata koeficijenta determinacije (R^2). Utvrđeno je da su optimalne vrednosti parametra k 12 za prvi podskup podataka i 25 za drugi podskup podataka, za obe godine.

Model veštačke neuronske mreže pokazao je najmanju srednju apsolutnu grešku (MAE) u 2018. godini, ali nije postigao najbolje rezultate u drugim metrikama. Da bismo identifikovali optimalne hiperparametre, koristili smo *Grid Search* metode unakrsne validacije iz *scikit-learn* biblioteke. Za obuku je korišćen Adam optimizator koji omogućava stabilnu konvergenciju i efikasno učenje. Za prvi podskup podataka, optimalan broj epoha bio je 20 u 2018. godini i 30 u 2019. godini, sa veličinama serija od 20, odnosno 10. Za drugi podskup podataka, optimalan broj epoha bio je 30 u 2018. i 50 u 2019. godini, sa veličinama serija od 10, odnosno 20. Broj korišćenih skrivenih neurona bio je 5 u svim skupovima podataka.

Dobijeni rezultati pokazuju da parametri obuke variraju u zavisnosti od podskupa podataka i godine, sa generalno dužim vremenom obuke (odnosno većim brojem epoha) potrebnim za podskup podataka za 2019. godinu u poređenju sa 2018. godinom. Ova varijacija se može pripisati razlikama u distribuciji podataka i složenosti između dve godine.

Numerički rezultati dobijeni u istraživanju prikazani su u tabeli 12 za prvi i tabeli 13 za drugi podskup podataka.

Tabela 12. Performanse regresionih modela za prvi podskup podataka.

Godina	Model	Trening metrike				Test metrike			
		MAE	MSE	RMSE	R ²	MAE	MSE	RMSE	R ²
2018	SVR	144	36486	191	0,11	138	33140	182	0,11
	RFR	25	1253	35	0,96	69	8295	91	0,78
	K-NN	69	10333	101	0,75	70	9491	97	0,75
	ANN	70	10353	101	0,75	65	8826	93	0,76
2019	SVR	167	45201	212	0,06	176	47096	217	0,07
	RFR	36	2272	47	0,95	95	14606	120	0,71
	K-NN	110	18574	136	0,61	117	20466	143	0,60
	ANN	117	20466	143	0,60	112	19137	138	0,62

Tabela 13. Performanse regresionih modela za drugi podskup podataka.

Godina	Model	Trening metrike				Test metrike			
		MAE	MSE	RMSE	R ²	MAE	MSE	RMSE	R ²
2018	SVR	113	26026	161	0,83	112	24037	155	0,83
	RFR	31	2024	44	0,98	81	13036	114	0,91
	K-NN	90	15637	125	0,90	91	15783	125	0,89
	ANN	99	19517	139	0,87	98	18662	136	0,87
2019	SVR	157	42586	206	0,75	160	43607	208	0,75
	RFR	32	1880	43	0,98	86	12918	113	0,92
	K-NN	104	18255	135	0,89	110	20392	142	0,88
	ANN	128	25854	160	0,85	131	26870	163	0,84

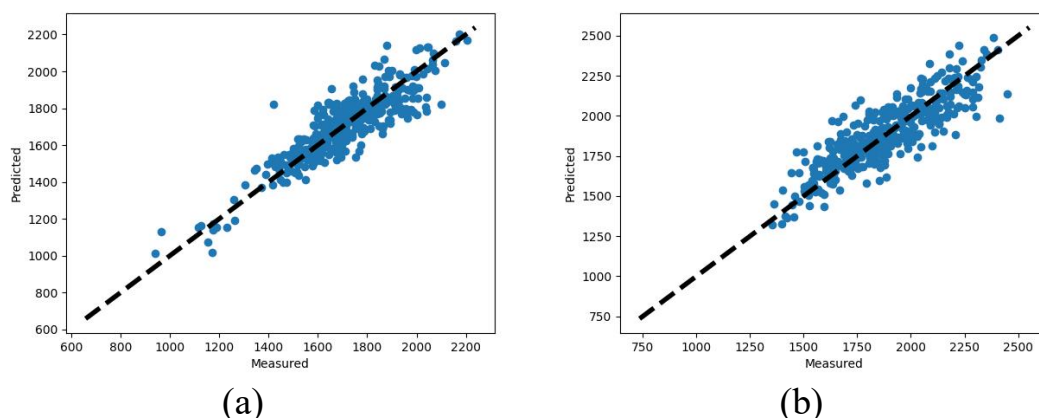
Eksperimenti su potvrdili da je RFR ostvario najprecizniju predikciju prinosa suncokretovog ulja u obe posmatrane godine, sa najvišim vrednostima koeficijenta determinacije (R² od 0,96 na trening skupu i 0,78 na test skupu za 2018. godinu, odnosno 0,95 i 0,71 za trening i test skupove za 2019. godinu), kao i najmanjim greškama (RMSE od 35 i 91 za 2018. godinu, odnosno 47 i 120 za 2019. godinu). Modeli K-NN i ANN ostvarili su međusobno slične rezultate, dok je SVR pokazao znatno lošije performanse u svim metrikama, što ukazuje na njegovu ograničenu primenljivost u ovom kontekstu.

Dobijeni rezultati demonstriraju robusnost metoda zasnovanih na ansamblu (kao što je RFR) za predviđanje prinosa suncokretovog ulja. Takođe, rezultati pokazuju i potencijal neuronskih mreža kada su podešene na odgovarajući način.

5.2.2. Vizuelna analiza predikcija

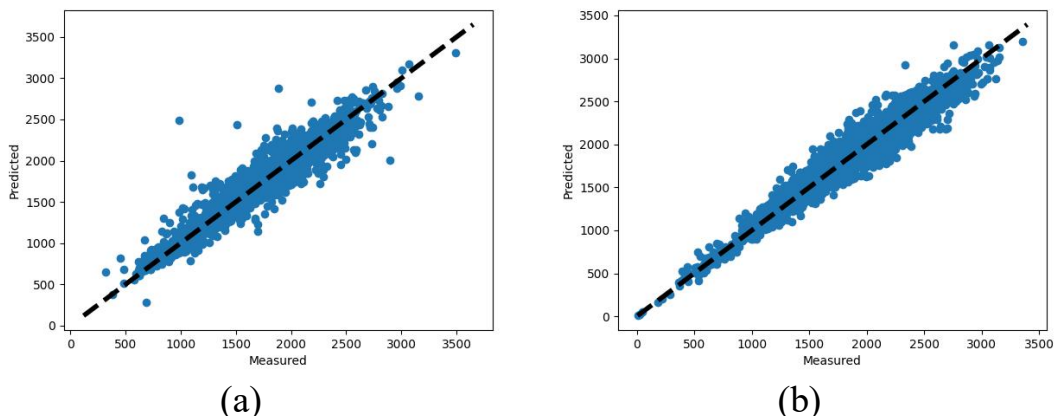
U cilju vizuelne analize tačnosti predikcije RFR modela, prikazani su odnosi između izmerenih i predviđenih vrednosti visine prinosa ulja za svaki od podskupova podataka za 2018. i 2019. godinu.

Na slici 24, prikazani su grafikoni na kojima svaka tačka predstavlja par stvarne i predviđene vrednosti, dok isprekidana linija označava idealno poklapanje ($y=x$), što omogućava jednostavno uočavanje odstupanja modela od stvarnih rezultata.



Slika 24. Izmerene i predviđene vrednosti za RFR model za prvi podskup podataka u 2018. (a) i 2019. godini (b).

Na slici 25 nalaze se grafikoni sa izmerenim i predviđenim vrednostima RFR algoritma za drugi korišćeni podskup podataka.

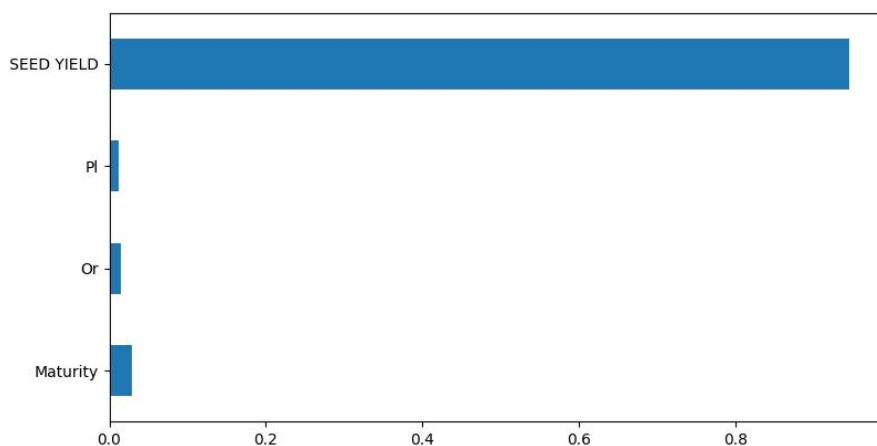


Slika 25. Izmerene i predviđene vrednosti za RFR model za drugi podskup podataka u 2018. (a) i 2019. godini (b).

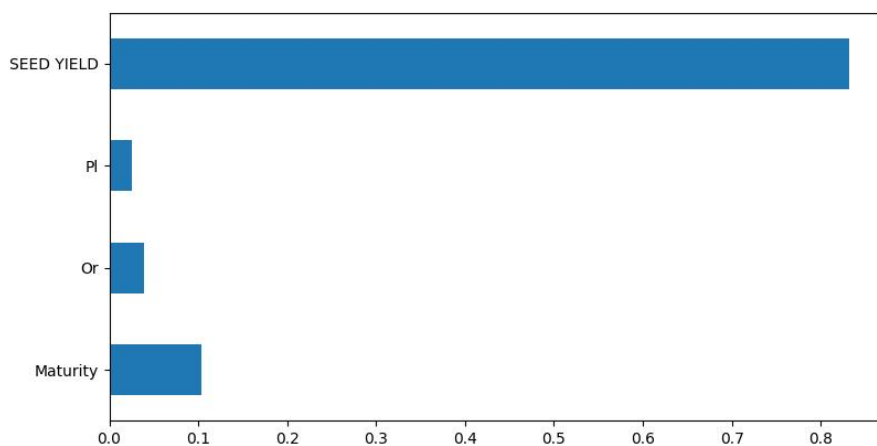
5.2.3. Analiza važnosti karakteristika

Važnost karakteristika (engl. *feature importance*) određena je korišćenjem ugrađene metrike algoritma slučajne šume. Ovaj metod procenjuje doprinos svake ulazne varijable u procesu donošenja odluka modela, obezbeđujući transparentnu procenu njihove prediktivne moći. Analiza je sprovedena korišćenjem svojstva *feature_importances_* iz obučenog RFR modela, koje svakoj ulaznoj karakteristici dodeljuje relativne ocene značaja u procesu donošenja odluke.

Kod prvog podskupa podataka utvrđeno je da su za obe godine najrelevantnije karakteristike koje utiču na predviđanje prinosa suncokretovog ulja prinos semena, otpornost na plamenjaču (Pl), otpornost na volovod (Or) i zrelost (slike 26 i 27). Ove karakteristike su pokazale jaku korelaciju sa visinom prinosa ulja, što ukazuje na njihovu ključnu ulogu u modelima predviđanja. Iste karakteristike su ostale važne i za 2018. i za 2019. godinu, što potvrđuje stabilnost rezultata, odnosno stabilnost modela i pouzdanost identifikovanih karakteristika.

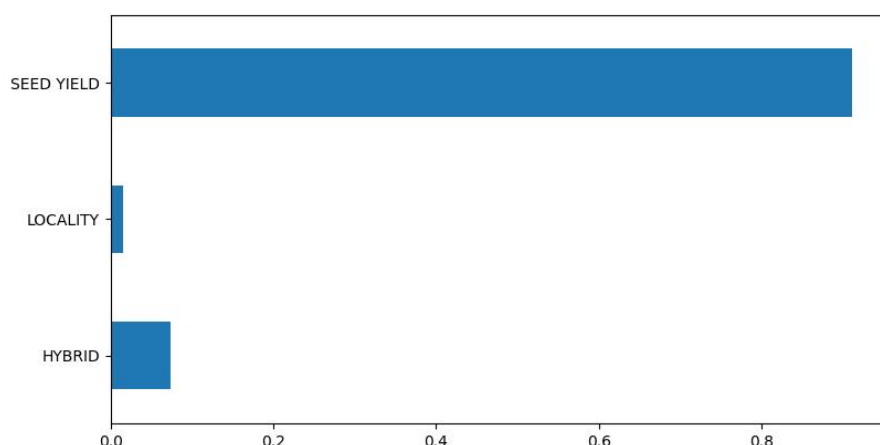


Slika 26. Najrelevantnije karakteristike za RFR u 2018. godini za prvi podskup podataka.

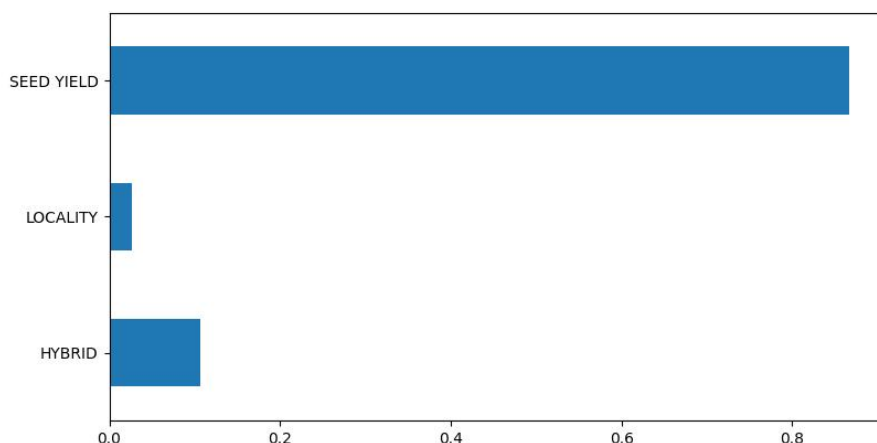


Slika 27. Najrelevantnije karakteristike za RFR u 2019. godini za prvi podskup podataka.

U drugom podskupu podataka je za predviđanje visine prinosa ulja, pored prinosa semena bila značajna i karakteristika koja predstavlja specifičan hibrid korišćen u studiji (slike 28 i 29). Isti trend je uočen tokom obe godine, što ukazuje da genetski sastav i specifične osobine svakog hibrida značajno utiču na prinos ulja. Ovaj rezultat je u skladu sa agronomskim uvidima, gde različiti hibridi suncokreta pokazuju različite obrasce rasta, akumulacije ulja i otpornost na stresore iz okoline.



Slika 28. Najrelevantnije karakteristike za RFR u 2018. godini za drugi podskup podataka.



Slika 29. Najrelevantnije karakteristike za RFR u 2019. godini za drugi podskup podataka.

Analiza važnosti karakteristika pruža uvid u karakteristike koje najviše doprinose u predikciji, što može poboljšati fino podešavanje modela. Uvidi dobijeni analizom važnosti karakteristika takođe mogu pomoći agronomima i farmerima u praksi, na primer pri odabiru hibrida visokog prinosa i optimizaciji uslova uzgoja.

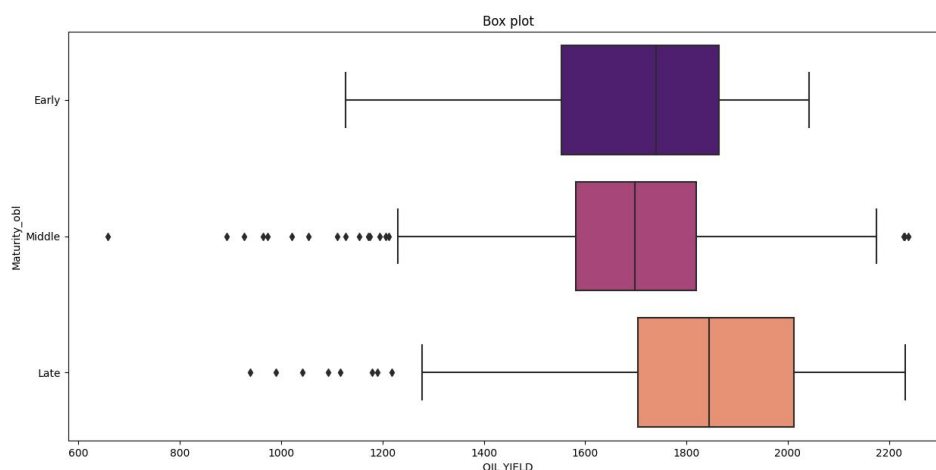
5.2.4 Statistička analiza i uticaj karakteristika na prinos ulja

Statistička analiza ima ključnu ulogu u identifikaciji odnosa između različitih karakteristika koje utiču na prinos suncokretovog ulja. Analizom podataka mogu se identifikovati obrasci i korelacije koje pružaju uvid u složene interakcije između uslova životne sredine, hibridnih karakteristika i visine prinosa ulja.

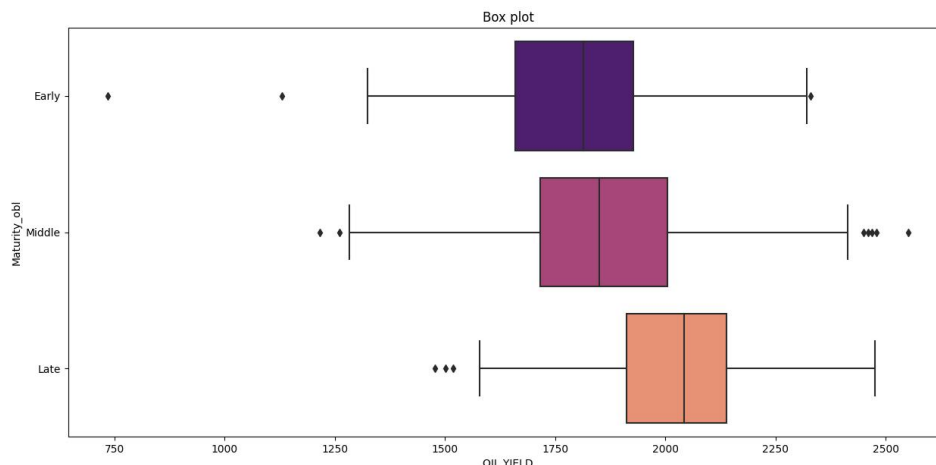
Jedan od doprinosa ove disertacije jeste identifikovanje veza između prinosa ulja, klimatskih faktora i genetičkih karakteristika hibrida suncokreta. Razumevanje ovih veza je od suštinskog značaja za optimizaciju selekcije hibrida, poboljšanje tehnika kultivacije i povećanje tačnosti prediktivnog modeliranja. Slike 30-35 pokazuju varijacije u prinosu ulja na koje utiču najznačajnije karakteristike identifikovane u našoj analizi. Ovi uvidi potvrđuju važnost odabira karakteristika u prediktivnom modeliranju i nude praktične implikacije za agronomsko donošenje odluka i programe oplemenjivanja hibrida koji imaju za cilj unapređenje tačnosti predikcije prinosa ulja.

Kao što je prikazano na slikama 30 i 31, zrelost hibrida je imala značajan uticaj na predviđanje visine prinosa suncokretovog ulja. Analiza podataka je pokazala da su hibridi kasnog zrenja nadmašili hibride srednjeg i ranog zrenja u pogledu visine prinosa ulja. Ovaj rezultat je u skladu sa očekivanjima, pošto hibridi koji kasnije sazrevaju generalno imaju duži ciklus rasta, što omogućava veću akumulaciju biomase i stvaranje ulja.

Međutim, uočeno je zanimljivo odstupanje u vezi sa hibridima ranog zrenja u 2018. godini. Za razliku od rezultata iz 2019. godine, rani hibridi u 2018. su pokazali širi opseg prinosa, što sugeriše da neki od ovih hibrida poseduju visok genetski potencijal za prinos ulja. Može se pretpostaviti da bi dalje istraživanje genetički indukovanih osobina ranih hibrida bilo korisno jer odabir i optimizacija takvih hibrida može dovesti do poboljšanih rezultata prinosa, posebno u regionima sa kraćim sezonama rasta ili nepovoljnim klimatskim uslovima.

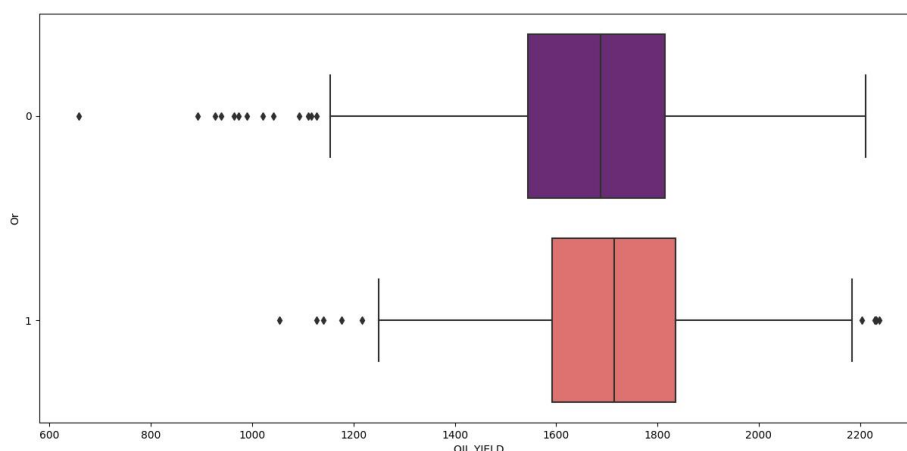


Slika 30. Vrednosti visine prinosa ulja u odnosu na karakteristiku zrelost hibrida u 2018. godini.



Slika 31. Vrednosti visine prinosa ulja u odnosu na karakteristiku zrelost hibrida u 2019. godini.

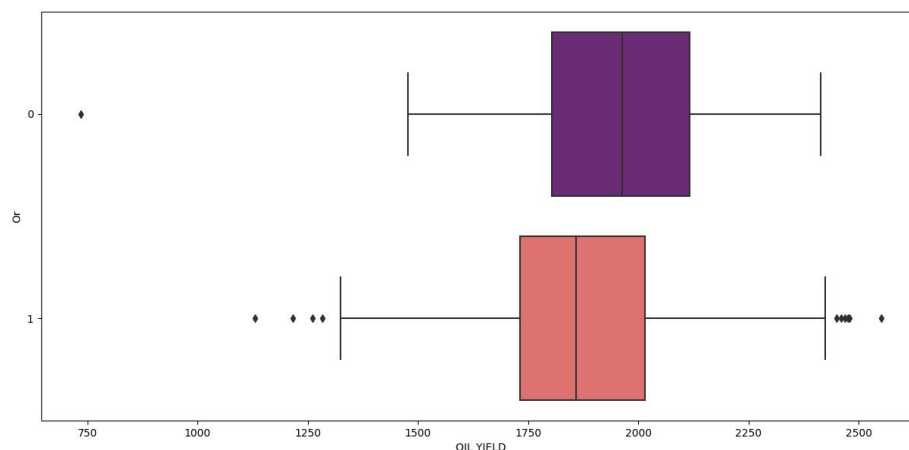
Na slici 32 se vidi da prisustvo otpornosti na volovod ima pozitivan uticaj na visinu prinosa ulja kod suncokreta u 2018. godini. Genotipovi koji poseduju ovo obeležje ostvarili su viši prinos ulja u poređenju sa genotipovima bez njega.



Slika 32. Vrednosti visine prinosa ulja u odnosu na otpornost na volovod u 2018. godini.

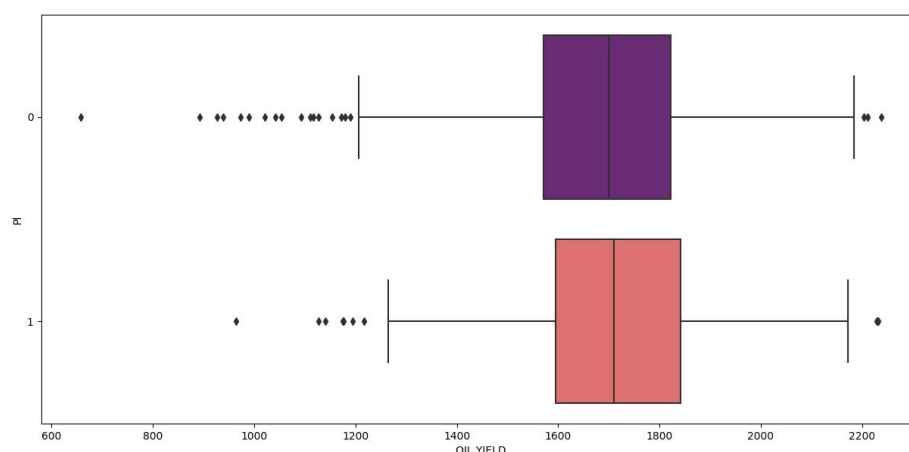
Na slici 33 koja prikazuje odnos visine prinosa ulja suncokreta u odnosu na Or obeležje u 2019. godini, vide se drugačiji rezultati u odnosu na 2018. godinu. Genotipovi koji ne poseduju otpornost na Or ostvarili su viši prinos ulja u poređenju sa otpornim genotipovima. U 2019. godini, blagi temperaturni stres u ranim fazama rasta (april–maj) i veća količina padavina tokom vegetacione sezone omogućili su neotpornim genotipovima da ostvare veći prinos ulja, dok su otporni genotipovi imali sporiji početak rasta zbog energetskih troškova povezanih sa mehanizmima za odbranu od parazita. Ovo

je u skladu sa konceptom *fitness cost* otpornosti, što znači da genotipovi sa genima za otpornost u odsustvu parazita mogu imati smanjenu produktivnost (Bergelson & Purrington, 1996; Strauss et al., 2002; Tian et al., 2003). Nasuprot tome, u 2018. godini, topliji početak sezone i umereno sušniji uslovi povećali su pritisak volovoda, što je dalo prednost otpornim genotipovima.

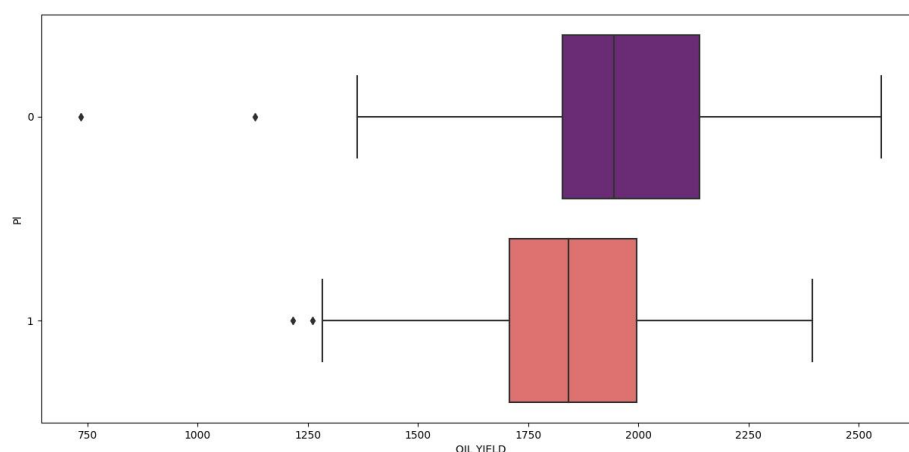


Slika 33. Vrednosti visine prinosa ulja u odnosu na otpornost na volovod u 2019. godini.

Slično kao kod otpornosti na volovod, genotipovi sa genima za otpornost na plamenjaču pokazali su varijabilnu povezanost sa prinosom ulja u zavisnosti od godine u kojoj je izvršeno merenje. U 2018. godini, kombinacija toplijeg početka vegetacije i umerenog sušnog stresa verovatno je uticala na povećanje efikasnosti otpornih genotipova, što se može videti na slici 34. Nasuprot tome, u 2019. godini su hladniji uslovi i veća količina padavina omogućili neotpornim sortama da ostvare veći prinos (slika 35). Ovi rezultati ukazuju na postojanje *genotip* × *okolina interakcije*, gde otpornost na oba patogena može biti povezana sa trošakom otpornosti (engl. *fitness cost*) u odsustvu jakog pritiska bolesti.



Slika 34. Vrednosti visine prinosa ulja u odnosu na otpornost na plamenjaču u 2018. godini.



Slika 35. Vrednosti visine prinosa ulja u odnosu na otpornost na plamenjaču u 2019. godini.

Otpornost na volovod i plamenjaču su se pokazale kao važne karakteristike za visinu prinosa suncokretovog ulja. Otporni hibridi su imali veći prinos ulja u 2018. godini dok su osjetljivi imali veći prinos u 2019. godini. Uvođenjem gena otpornosti u hibride suncokreta, kao što su Or i Pl, često se uvode i druge osobine koje utiču na smanjenje prinosa semena i ulja (Dedić et al., 2022). Potrebno je nekoliko ciklusa ukrštanja da bi se eliminisale negativne karakteristike. Materijal iz ranijih ciklusa oplemenjivanja potrebno je dodatno unaprediti u pogledu prinosa ulja.

Analiza odabranih hibrida pokazala je da su hibridi dobijeni od kasno zrele linije DF-AB-2 imali najveći prinos ulja u 2018. godini (tabela 14), dok su hibridi zasnovani na srednje sazreloj liniji HA-26 ostvarili najbolje rezultate u 2019. godini (tabela 15).

Tabela 14. Hibridi sa najvećim prinosom ulja za 2018. godinu.

Hibrid	Prinos ulja	Or	Pl	Zrelost	Opis
998	2238,06	1	0	Kasni	Hibrid na bazi majke DF-AB-2
363	2231,15	1	1	Kasni	Hibrid na bazi majke DF-AB-2
773	2230,73	1	1	Srednji	Kontrola
798	2228,16	1	1	Srednji	Kontrola
139	2211,46	0	0	Kasni	Hibrid na bazi majke DF-AB-2
792	2203,33	1	0	Kasni	Hibrid na bazi majke KINA-B-5
162	2183,30	1	0	Kasni	Hibrid na bazi majke DF-AB-2
372	2179,68	0	0	Kasni	Hibrid na bazi majke DF-AB-2
890	2175,02	0	0	Kasni	Hibrid na bazi majke KINA-B-5
760	2172,36	1	0	Srednji	Hibrid na bazi majke IMI-AB-12

Tabela 15. Hibridi sa najvećim prinosom ulja za 2019. godinu.

Hibrid	Prinos ulja	Or	Pl	Zrelost	Opis
848	2551,85	1	0	Srednji	Hibrid na bazi majke HA-26
116	2480,18	1	0	Srednji	Hibrid na bazi majke HA-26
550	2475,18	1	0	Kasni	Hibrid na bazi majke DF-AB-2
114	2469,05	1	0	Srednji	Hibrid na bazi majke HA-26
773	2460,96	1	0	Srednji	Kontrola
112	2450,10	1	0	Srednji	Hibrid na bazi majke HA-26
350	2424,00	1	0	Kasni	Hibrid na bazi majke DF-AB-2
229	2414,04	0	0	Srednji	Hibrid na bazi majke HA-26
325	2403,94	1	0	Kasni	Hibrid na bazi majke DF-AB-2
873	2395,66	1	0	Srednji	Kontrola

6. Diskusija

U okviru ovog poglavlja analizirani su rezultati eksperimenata koji su predstavljeni u prethodnom poglavlju.

6.1 Diskusija rezultata za detekciju stresa usled suše

Uprkos malom i klasno neravnomerno raspoređenom skupu podataka, modeli neuronskih mreža DenseNet, Inception i MobileNet su uspešno prepoznali biljke koje su bile pod stresom usled suše, sa tačnošću u rasponu 0,85–0,93 i F1-skorom između 0,87–0,93.

MobileNet se izdvojio kao najefikasniji i najstabilniji model, kako u pogledu performansi, tako i u pogledu brzine obučavanja. Iako su DenseNet i Inception takođe dali dobre rezultate, njihova doslednost u performansama bila je slabija, naročito kada su analizirani različiti formati podataka, a uz to su zahtevali i duže vreme obuke. Ova činjenica sugerise da, iako duboke neuronske mreže sa složenijom arhitekturom mogu postići visoku tačnost, njihova primena u poljoprivrednim sistemima mora biti pažljivo razmotrena u kontekstu efikasnosti i dostupnih resursa (Sabo et al., 2023).

Kada je u pitanju primena augmentacije podataka, pokazalo se da standardna augmentacija nije doprinela poboljšanju performansi modela, dok je ciljana augmentacija donela vidljive koristi. Konkretno, primena ciljane augmentacije na slikama u sivom tonu pokazala se kao posebno efikasna, rezultujući poboljšanjem performansi, pri čemu su za MobileNet ostvareni F1-skor i tačnost od 0,95. Ovaj rezultat može se objasniti time što je ciljano proširivanje podataka omogućilo bolju diferencijaciju karakteristika između klasa, što je omogućilo modelu njihovo efikasnije razlikovanje. Ovi rezultati ukazuju na to da primena ciljane augmentacije može predstavljati jedno od potencijalnih rešenja za problem ograničenih podataka, što je značajno u razvoju AI sistema koji se primenjuju u poljoprivredi.

Još jedan rezultat ovog istraživanja jeste to da su slike u sivom tonu dale bolje rezultate u poređenju sa originalnim slikama u boji, naročito kada su kombinovane sa ciljanim metodama augmentacije. Ovo može biti posledica smanjene složenosti podataka u sivom formatu, što omogućava modelu da se bolje fokusira na strukturne karakteristike biljke bez uticaja varijacija u boji, koje ponekad mogu dodati suvišne informacije i otežati proces klasifikacije. Ovakav pristup može biti naročito koristan kada se modeli koriste u aplikacijama sa ograničenim računarskim resursima, jer smanjuje potrebu za obradom podataka i time omogućava brže izvođenje predikcija.

Korišćena metodologija generiše binarne rezultate, što znači da predikcije ukazuju samo na to da li je biljka izložena stresu usled suše ili ne. Ovakav jednostavan izlaz modela olakšava interpretaciju rezultata stručnjacima u poljoprivredi i omogućava brzo donošenje odluka. Posebno je značajno što se ovakav pristup može koristiti za određivanje inbred linija koje, kada se uzgajaju u kontrolisanim uslovima, pokazuju nižu toleranciju na stres usled suše. Na taj način, dobijeni rezultati mogu služiti kao osnova za donošenje informisanih odluka o izboru inbred linija koje su otpornije i bolje prilagođene specifičnim ekološkim uslovima.

Kada je reč o praktičnoj primeni u poljoprivredi, identifikacija otpornijih linija može imati važnu ulogu u optimizaciji upravljanja poljoprivrednim resursima. Konkretno, informacije dobijene ovim modelom mogu pomoći pri donošenju odluka o navodnjavanju, što je posebno značajno u proizvodnji semena. Na osnovu ovih rezultata vodni resursi bi mogli efikasnije da se rasporede, a dodatna nega da se pruži biljkama koje su manje tolerantne na sušu. S obzirom na rastuće izazove povezane sa klimatskim promenama i sve veću potrebu za preciznom poljoprivredom, ovakav pristup može doprineti održivijem i produktivnijem upravljanju usevima.

Rezultati ovog istraživanja u skladu su sa nalazima prethodnog rada autora Penmetsa & Sarma (2022) koji su, koristeći transferno učenje i MobileNet, pokazali efikasnost modela za male, neuravnotežene skupove poljoprivrednih podataka. U istraživanju je postignuta tačnost od 89,5% uz pomoć standardnih RGB slika. U okviru ovog istraživanja, primena ciljane augmentacije nad slikama u sivim tonovima rezultovala je visokim performansama (F1-skor od 0,95). To ukazuje na mogućnost da kombinacija ciljane augmentacije i redukcije kolornih informacija doprinosi izdvajanju strukturnih karakteristika povezanih sa stresom usled suše.

Slični zaključci uočeni su i kod autora Patra & Sahoo (2024) čiji je fokus bio na objašnjivost korišćenog modela neuronske mreže. Njihov MobileNetV3-GradCAM pristup postigao je tačnost od 94,2%, uz skup podataka od oko 1500 slika, što je značajno veće od skupa podataka iz ovog istraživanja. U njihovom radu, Grad-CAM mape modela dosledno su isticale margine listova i teksturu zemljišta kao ključne indikatore suše, što je u skladu sa rezultatima dobijenim u ovom istraživanju. Dobijeni rezultati ukazuju na to da obrada slika u sivim tonovima može doprineti efikasnijem isticanju strukturnih karakteristika biljke.

Obe studije potvrđuju da je MobileNet pogodan za primenu u kontekstu poljoprivrednih podataka u cilju detekcije stresa usled suše. Dobijeni rezultati

takođe ukazuju na potencijal primene ciljane augmentacije i obrade slika u sivim tonovima u situacijama sa ograničenim skupovima podataka.

Buduća istraživanja mogla bi biti usmerena ka integraciji metoda prethodne obrade korišćenih u ovom istraživanju sa tehnikama interpretabilnosti, poput Grad-CAM pristupa, kako bi se razvili modeli koji istovremeno ostvaruju visoku tačnost i omogućavaju jasniju interpretaciju procesa donošenja odluka.

6.2 Diskusija rezultata za predviđanje visine prinosa ulja

6.2.1 Analiza karakteristika hibrida sa najvećim prinosom ulja

Među 1250 kombinacija hibrida suncokreta testiranih tokom dve godine ispitivanja, specifični hibridi su pokazali najveće vrednosti prinosa ulja, što je dovelo do njihovog izbora za dalju analizu. Hibridi suncokreta razvijeni su kontrolisanim ukrštanjem majčinskih i očevih linija, pri čemu svaki roditelj doprinosi genetičkim osobinama koje utiču na sadržaj ulja, otpornost na bolesti i ukupnu prilagodljivost. Majčinska linija, međutim, igra posebno značajnu ulogu u oblikovanju hibridnih performansi zbog nasleđivanja njene celokupne citoplazme, što utiče na ključne fiziološke i biohemijske procese (Camus et al., 2022). Imajući u vidu ovaj uticaj majke, hibridi se mogu klasifikovati na osnovu karakteristika njihovih matičnih linija, što omogućava strukturiraniju procenu potencijala prinosa.

Analizom odabranih hibrida utvrđeno je da su dve matične linije preovlađivale među kombinacijama sa najvećim prinosom: kasno zrela linija DF-AB-2 iz 2018. godine i srednje zrela HA-26 iz 2019. godine. Hibridi dobijeni od DF-AB-2 su pokazali superiorni prinos ulja u 2018. godini, koju karakterišu povišene temperature i manje padavine. Ovo je u skladu sa shvatanjem da hibridi sa dužim sezonama rasta, kao što su oni dobijeni od DF-AB-2, pokazuju veću otpornost na periode suše i visoke temperature (Ćuk et al., 2022). Produžene vegetativne i reproduktivne faze omogućavaju ovim hibridima da optimizuju unos vode i hranljivih materija tokom dužeg perioda, ublažavajući štetne efekte sušnih intervala.

S druge strane, hibridi razvijeni iz matične linije HA-26 imali su bolji učinak u 2019. godini, godini sa povoljnijom distribucijom padavina. Uspeh hibrida zasnovanih na HA-26 u ovoj određenoj godini sugerise da srednje zrele linije mogu biti bolje prilagođene sredinama gde je dostupnost vode dovoljna, ali zahteva efikasno korišćenje tokom celog ciklusa rasta (Škorić et al., 2004). Ovi nalazi naglašavaju važnost pažljivog odabira matičnih linija na osnovu

uslova životne sredine i specifičnih klimatskih izazova koji se očekuju tokom sezone rasta.

Razumevanje uticaja majčinog nasleđa na prinos suncokretovog ulja pruža vredan uvid u programe oplemenjivanja hibrida. Odabirom matičnih linija sa osobinama koje su u skladu sa specifičnim klimatskim uslovima, uzgajivači mogu razviti hibride koji dosledno daju visoke prinose u različitim scenarijima životne sredine. Dalja istraživanja genetičkih i fizioloških mehanizama koji stoje iza citoplazmatskog nasleđa mogla bi pomoći da se preciziraju strategije hibridne selekcije, što bi dovelo do otpornijih i produktivnijih sorti suncokreta u budućnosti.

6.2.2 Diskusija primene metoda mašinskog učenja za predviđanje prinosa ulja

Rezultati dobijeni primenom različitih algoritama mašinskog učenja pokazali su da izbor algoritma utiče na uspešnost predikcije visine prinosa ulja. Model koji je ostvario najbolje performanse je model koji koristi RFR algoritam. On je ostvario najbolje performanse za obe eksperimentalne godine, oba podskupa podataka i većinu korišćenih evalucionih metrika.

Potencijalni razlog zbog koga je RFR ostvario najbolje performanse je to što RFR uglavnom dobro funkcioniše kada postoji veći broj potencijalno relevantnih karakteristika koje nisu međusobno linearno zavisne. Izračunavanje visine prinosa ulja predstavlja složen proces na koji utiče veliki broj faktora, kao što su genetičke karakteristike hibrida, prinos semena, otpornost na bolesti i parazite, i uslovi sredine. Stabilni rezultati u obe eksperimentalne godine pokazuju robusnost modela u različitim eksperimentalnim uslovima.

K-NN je ostvario rezultate odmah iza RFR modela što ukazuje na to da se i ovaj algoritam može koristiti za predikciju visine prinosa ulja. Pošto K-NN koristi sličnost karakteristika hibrida za predikciju, to znači da se obrasci među sličnim hibrida mogu koristiti za predikciju visine prinosa ulja. Međutim, K-NN zavisi od izbora parametra k i osetljiviji je na distribuciju podataka, zbog čega ne može direktno da uči kompleksne odnose kao RFR.

ANN model je pokazao zadovoljavajuće performanse, ali nije nadmašio RFR. Mogući uzrok ovakvih rezultata može biti relativno mali skup podataka u odnosu na onaj koji je potreban za veći potencijal neuronskih mreža. Podešavanjem hiperparametara dobijeno je stabilno obučavanje modela, te rezultati potvrđuju potencijal neuronskih mreža za primenu u agronomiji kod predikcije prinosa ulja.

Model koji je pokazao najslabije performanse je SVR. Mogući razlozi za ove rezultate su to što ovaj model ima ograničenu sposobnost da uhvati veliki broj interakcija bez odgovarajuće transformacije podataka i što izbor parametara može značajno da utiče na rezultat.

Spoljašnji faktori poput temperature i količine padavina uticali su na stresne uslove tokom vegetacije, a oni su se razlikovali u 2018. i 2019. godini. Iako ovi faktori utiču na visinu prinosa ulja kod suncokreta, RFR model je zadržao stabilne performanse za predikciju. Ovo pokazuje upotrebljivost modela za primenu u promenljivim spoljašnjim uslovima.

Karakteristike koje su se pokazale kao najvažnije za predikciju visine prinosa ulja su prinos semena, hibrid, zrelost hibrida, kao i njihova otpornost na volovod i plamenjaču. Prinos semena se izdvojio kao najvažnija karakteristika, što je očekivano s obzirom na njegovu direktnu povezanost sa količinom formiranog ulja. Pored toga, hibrid je takođe izdvojen kao značajna karakteristika jer različiti hibridi imaju različit potencijal za akumulaciju ulja, adaptaciju i odgovor na stres. Takođe, zrelost hibrida u odnosu na vegetacionu sezonu može da utiče na konačni prinos, a otpornost hibrida na parazite kao što su volovod i plamenjača očekivano poboljšava prinos.

Dobijeni rezultati potvrđuju da ansambl modeli poput RFR imaju najveći potencijal za primenu u predikciji prinosa suncokretovog ulja. Buduća istraživanja mogla bi se bazirati na proširenju skupova podataka uključivanjem dodatnih lokacija i eksperimentalnih sezona kako bi se proverila generalizacija RFR modela sa novim podacima. Dodatno, mogli bi se uključiti novi meteorološki podaci, kao što su satelitski vegetacioni indeksi, koji bi doprineli boljem opisu uslova životne sredine. Kada je reč o metodološkom unapređenju istraživanja, planira se ispitivanje naprednih ansambl metoda i dubokih modela prilagođenih tabularnim podacima, kao i razvoj hibridnih pristupa koji kombinuju različite tehnike mašinskog učenja. Radi boljeg razumevanja dobijenih rezultata i primene rezultata u realnim sistemima, planirano je i uključivanje objašnjive veštačke inteligencije.

7. Zaključak

Doktorska disertacija bila je usmerena na primenu savremenih metoda veštačke inteligencije za pravovremenu detekciju stresa usled suše i predviđanje visine prinosa ulja kod suncokreta. Istraživanjem su ispitane mogućnosti primene ovih metoda u cilju razvoja održive poljoprivrede kroz efikasniju upotrebu resursa. Korišćene su metode mašinskog i dubokog učenja na tabularnim i slikovnim podacima kako bi se unapredilo razumevanje faktora koji utiču na produktivnost i otpornost suncokreta u promenljivim agroekološkim uslovima.

Istraživanje je pokazalo da se veštačka inteligencija može uspešno primeniti za detekciju strasa usled suše i predviđanje visine prinosa ulja kod biljaka suncokreta. U kontekstu detekcije stresa usled suše, konvolucione mreže su sa visokom preciznošću pokazale sposobnost prepoznavanja vizuelnih indikatora stresa na biljkama suncokreta. Konkretno, MobileNet je ostvario 0,95 za F1-skor i tačnost. Kada je reč o predikciji visine prinosa suncokretovog ulja, model koji je pokazao najbolje performanse bio je RFR model. On je ostvario prosečno 0,83 za R^2 vrednost za dva podskupa podataka u dve eksperimentalne sezone.

Rezultati istraživanja potvrdili su predložene hipoteze.

Hipoteze H1 i H2 potvrđene su kroz uspešnu primenu modela dubokog učenja za detekciju stresa usled suše kod suncokreta. Rezultati su pokazali da slike korena i nadzemnih organa suncokreta sadrže relevantne informacije o morfološkim karakteristikama povezanim sa promenama usled nedostatka vode. Ovo je omogućilo razlikovanje biljaka izloženih stresu od kontrolnih biljaka, čime je potvrđeno da vizuelni fenotip suncokreta predstavlja relevantan izvor informacija za detekciju stresa usled suše primenom metoda dubokog učenja.

Hipoteza H3 je potvrđena rezultatima dobijenim korišćenjem modela sa augmentovanim podacima. Da se performanse modela dubokog učenja mogu poboljšati primenom odgovarajuće tehnike augmentacije podataka takođe je pokazano dobijenim rezultatima.

Primenom različitih algoritama mašinskog učenja za predikciju visine prinosa suncokretovog ulja na osnovu genetičkih karakteristika hibrida i agroekoloških podataka lokacija proizvodnje potvrđena je hipoteza H4. Najbolje performanse ostvarene su upotrebom RFR modela, sa prosečnom vrednošću koeficijenta determinacije R^2 od 0,83 za analizirane skupove podataka.

Hipoteza H5 je potvrđena analizom značaja ulaznih karakteristika i statističkom analizom uticaja pojedinačnih osobina. Takođe je pokazano da genetičke karakteristike hibrida predstavljaju važne faktore povezane sa ostvarivanjem visokog prinosa ulja i mogu da utiču na odabir odgovarajućih hibrida u različitim agroekološkim uslovima.

Najznačajnija ograničenja ovog istraživanja odnose se na veličinu i raznovrsnost skupova podataka, ograničenja u hardverskim resursima kao i ograničenu dostupnost genotipskih i fenotipskih podataka. Uprkos navedenim ograničenjima, postignuti rezultati ukazuju na značajan potencijal pristupa baziranog na veštačkoj inteligenciji za automatizaciju i optimizaciju procesa u proizvodnji suncokreta.

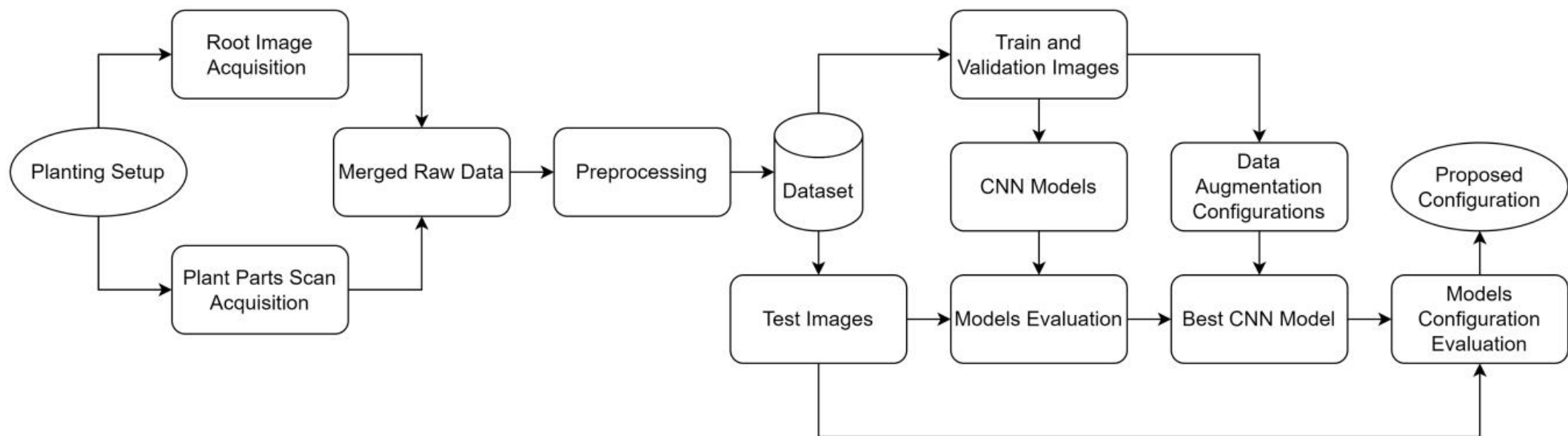
Istraživanje bi moglo da se nastavi u nekoliko različitih pravaca. Osnovni dalji pravac istraživanja orijentisan je ka proširenju skupova podataka, uključivanjem većeg broja genotipova, hibrida, sezona, kao i podataka iz drugih klimatskih regiona. U metodološkom smislu, bilo bi zanimljivo nastaviti istraživanje kroz primenu naprednijih arhitektura dubokog učenja, kao što su multimodalne mreže za kombinovanje vizuelnih i numeričkih informacija. Kada je reč o primeni u realnim sistemima, akcenat bi bio na razvoju interaktivnih sistema za podršku odlučivanju koji bi omogućili agronomima i istraživačima da u realnom vremenu procene stanje useva i optimizuju navodnjavanje i primenu đubriva. Ovi sistemi mogli bi onda biti integrisani u sisteme precizne poljoprivrede i mobilne aplikacije namenjene praktičnoj primeni u proizvodnji.

Na osnovu ovog istraživanja, može se zaključiti da su savremene metode veštačke inteligencije važan alat za budućnost agronomskih istraživanja. Primena modela veštačke inteligencije u oblasti suncokreta otvara mogućnosti za povećanje prinosa, optimizaciju potrošnje resursa i pravovremeno prepoznavanje stresa biljaka, čime se doprinosi razvoju održive i klimatski pametne poljoprivrede. Ova disertacija takođe može da predstavlja temelj za dalja istraživanja u oblasti digitalne agronomije, kao i za primenu sličnih metodoloških pristupa u drugim usevima i agroekološkim sistemima.

Dodatak

Tabela A.1. Prosečni meteorološki podaci za 2018. i 2019. godinu za sva četiri lokaliteta.

Lokalitet	Koordinate	Tip zemljišta	Godina	Prosečna temperatura (°C)							Odstupanje od višegodišnjeg proseka (°C)	Broj dana				Količina padavina u vegetacionom periodu	
				April	Maj	Jun	Jul	Avgust	Septembar	Ukupno		Tmax >20°C	Tmax >30°C	Tmax >35°C	Sa padavinama	U mm	U %
Subotica	46°05'53"N; 19°40'16"E	Peskovito zemljište	2018	16,5	20,4	21,6	23,0	24,3	18,0	20,63	+2,1	169	35	0	51	312	93
			2019	13,3	15,0	23,4	23,0	24,0	17,7	19,40	+1,2	131	31	1	56	431	129
Novi Sad	45°19'51"N; 19°50'59 E	Černozem	2018	17,2	20,4	21,4	21,9	24,0	18,5	20,57	+2,3	173	37	0	51	436	121
			2019	13,5	14,7	23,2	23,3	24,3	18,2	19,53	+1,4	126	40	1	52	415	115
Kikinda	45°43'11"N; 20°18'07"E	Humoglej i zaslanjena zemljišta	2018	16,7	20,6	21,4	22,4	24,2	18,6	20,65	+2,1	176	43	0	55	409	124
			2019	13,2	15,1	23,3	22,5	24,5	18,2	19,47	+1,3	133	38	3	55	399	121
Vršac	44°58'25"N, 21°13'17" E	Černozem	2018	17,3	20,3	20,9	22,1	24,1	18,4	20,52	+2,0	171	48	0	45	447	115
			2019	13,2	14,8	23,3	22,9	24,5	18,6	19,55	+1,0	129	34	2	44	304	78



Slika A.1. *Pipeline* predložene metodologije za detekciju stresa usled suše kod suncokreta (Lazić et al., 2024).

Reference

- Abbas, F. *et al.* (2020) “Crop Yield Prediction through Proximal Sensing and Machine Learning Algorithms,” *Agronomy*, 10(7), p. 1046. Available at: <https://doi.org/10.3390/agronomy10071046>.
- Abduljabbar, R. *et al.* (2019) “Applications of Artificial Intelligence in Transport: An Overview,” *Sustainability*, 11(1), p. 189. Available at: <https://doi.org/10.3390/su11010189>.
- Acquaah, G. (2012) *Principles of Plant Genetics and Breeding*. 1st ed. Wiley. Available at: <https://doi.org/10.1002/9781118313718>.
- Agrawal, R., Imieliński, T. and Swami, A. (1993) “Mining association rules between sets of items in large databases,” *Proceedings of the 1993 ACM SIGMOD international conference on Management of data. SIGMOD/PODS93: Joint ACM SIGMOD International Conference on Management of Data and ACM SIGMOD*, Washington D.C. USA: ACM, pp. 207–216. Available at: <https://doi.org/10.1145/170035.170072>.
- Ahmed, O. and Lindblad, O. (2022) *A Comparative Study on the Effect of Different Hyperparameters on the Performance of VGGNet-16 for Detection of Cardiomegaly in Chest X-ray Images*. Degree Project, First Cycle, 15 credits. KTH Royal Institute of Technology. Available at: <https://www.diva-portal.org/smash/get/diva2:1702465/FULLTEXT01.pdf> (Accessed: November 30, 2025).
- Al Sahili, Z. and Awad, M. (2022) “The power of transfer learning in agricultural applications: AgriNet,” *Frontiers in Plant Science*, 13, p. 992700. Available at: <https://doi.org/10.3389/fpls.2022.992700>.
- Ali, T. *et al.* (2024) “Smart agriculture: utilizing machine learning and deep learning for drought stress identification in crops,” *Scientific Reports*, 14(1), p. 30062. Available at: <https://doi.org/10.1038/s41598-024-74127-8>.
- Amankulova, K. *et al.* (2023) “Sunflower crop yield prediction by advanced statistical modeling using satellite-derived vegetation indices and crop phenology,” *Geocarto International*, 38(1), p. 2197509. Available at: <https://doi.org/10.1080/10106049.2023.2197509>.
- Amara, J., Bouaziz, B. and Algergawy, A. (2017) “A Deep Learning-based Approach for Banana Leaf Diseases Classification.” *Datenbanksysteme für Business, Technologie und Web (BTW 2017) - Workshopband*, Gesellschaft

für Informatik e.V., pp. 79–88. Available at: <https://dl.gi.de/items/13766147-8092-4f0a-b4e1-8a11a9046bdf> (Accessed: November 28, 2025).

An, J. *et al.* (2019) “Identification and Classification of Maize Drought Stress Using Deep Convolutional Neural Network,” *Symmetry*, 11(2), p. 256. Available at: <https://doi.org/10.3390/sym11020256>.

Andrianasolo, F.N. *et al.* (2014) “Prediction of sunflower grain oil concentration as a function of variety, crop management and environment using statistical models,” *European Journal of Agronomy*, 54, pp. 84–96. Available at: <https://doi.org/10.1016/j.eja.2013.12.002>.

Antonie, M.L., Zaiane, O.R. and Coman, A. (2001) “Application of data mining techniques for medical image classification,” *Proceedings of the Second International Conference on Multimedia Data Mining*, pp. 94–101.

Araújo, S.O. *et al.* (2021) “Characterising the Agriculture 4.0 Landscape—Emerging Trends, Challenges and Opportunities,” *Agronomy*, 11(4), p. 667. Available at: <https://doi.org/10.3390/agronomy11040667>.

Arik, S.Ö. *et al.* (2017) “Deep Voice: Real-Time Neural Text-to-Speech.” *Proceedings of the 34th International Conference on Machine Learning (ICML)*, Sydney, Australia: PMLR (Proceedings of Machine Learning Research), pp. 195–204. Available at: <https://proceedings.mlr.press/v70/arik17a.html>.

Arnell, N. (1999) “Climate change and global water resources,” *Global Environmental Change*, 9, pp. S31–S49. Available at: [https://doi.org/10.1016/S0959-3780\(99\)00017-5](https://doi.org/10.1016/S0959-3780(99)00017-5).

Azimi, S., Kaur, T. and Gandhi, T.K. (2021) “BAT Optimized CNN Model Identifies Water Stress in Chickpea Plant Shoot Images,” *2020 25th International Conference on Pattern Recognition (ICPR)*. *2020 25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy: IEEE, pp. 8500–8506. Available at: <https://doi.org/10.1109/ICPR48806.2021.9412720>.

Badshah, A. *et al.* (2024) “Crop Classification and Yield Prediction Using Robust Machine Learning Models for Agricultural Sustainability,” *IEEE Access*, 12, pp. 162799–162813. Available at: <https://doi.org/10.1109/ACCESS.2024.3486653>.

Barradas, A. *et al.* (2021) “Comparing Machine Learning Methods for Classifying Plant Drought Stress from Leaf Reflectance Spectra in *Arabidopsis thaliana*,” *Applied Sciences*, 11(14), p. 6392. Available at: <https://doi.org/10.3390/app11146392>.

Bergelson, J. and Purrington, C.B. (1996) “Surveying Patterns in the Cost of Resistance in Plants,” *The American Naturalist*, 148(3), pp. 536–558. Available at: <https://doi.org/10.1086/285938>.

Bishop, C.M. (2006) *Pattern recognition and machine learning*. New York: Springer (Information science and statistics).

Burkov, A. (2019) *The hundred-page machine learning book*. Polen: Andriy Burkov.

Călin, A.D., Coroiu, A.M. and Mureșan, H.B. (2023) “Analysis of Preprocessing Techniques for Missing Data in the Prediction of Sunflower Yield in Response to the Effects of Climate Change,” *Applied Sciences*, 13(13), p. 7415. Available at: <https://doi.org/10.3390/app13137415>.

Campbell, J.B., Wynne, R.H. and Thomas, V.A. (2023) *Introduction to remote sensing*. 6. edition. New York: Guilford Press.

Campbell, M., Hoane, A.J. and Hsu, F. (2002) “Deep Blue,” *Artificial Intelligence*, 134(1–2), pp. 57–83. Available at: [https://doi.org/10.1016/S0004-3702\(01\)00129-1](https://doi.org/10.1016/S0004-3702(01)00129-1).

Camus, M.F. *et al.* (2022) “Inheritance through the cytoplasm,” *Heredity*, 129(1), pp. 31–43. Available at: <https://doi.org/10.1038/s41437-022-00540-2>.

Cao, J. *et al.* (2020) “Identifying the Contributions of Multi-Source Data for Winter Wheat Yield Prediction in China,” *Remote Sensing*, 12(5), p. 750. Available at: <https://doi.org/10.3390/rs12050750>.

Çetin, N. *et al.* (2021) “Comparative Evaluation of Some Quality Characteristics of Sunflower Oilseeds (*Helianthus annuus* L.) Through Machine Learning Classifiers,” *Food Analytical Methods*, 14(8), pp. 1666–1681. Available at: <https://doi.org/10.1007/s12161-021-02002-7>.

Chandel, N.S. *et al.* (2021) “Identifying crop water stress using deep learning models,” *Neural Computing and Applications*, 33(10), pp. 5353–5367. Available at: <https://doi.org/10.1007/s00521-020-05325-4>.

Chandel, N.S. *et al.* (2022) “Water Stress Identification of Winter Wheat Crop with State-of-the-Art AI Techniques and High-Resolution Thermal-RGB Imagery,” *Plants*, 11(23), p. 3344. Available at: <https://doi.org/10.3390/plants11233344>.

Chandel, N.S. *et al.* (2024) “State-of-the-art AI-enabled mobile device for real-time water stress detection of field crops,” *Engineering Applications of*

Artificial Intelligence, 131, p. 107863. Available at: <https://doi.org/10.1016/j.engappai.2024.107863>.

Chapelle, O., Scholkopf, B., & Zien, A. (2009). Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]. *IEEE Transactions on Neural Networks*, 20(3), 542-542.

Chen, W. *et al.* (2024) “Artificial Intelligence in Logistics Optimization with Sustainable Criteria: A Review,” *Sustainability*, 16(21), p. 9145. Available at: <https://doi.org/10.3390/su16219145>.

Chollet, F. (2015) keras, GitHub. <https://github.com/fchollet/keras>

Concepcion Ii, R. *et al.* (2020) “Lettuce Leaf Water Stress Estimation Based on Thermo-Visible Signatures Using Recurrent Neural Network Optimized by Evolutionary Strategy,” *2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC)*. *2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC)*, Kuching, Malaysia: IEEE, pp. 1–6. Available at: <https://doi.org/10.1109/R10-HTC49770.2020.9356963>.

Coulibaly, S. *et al.* (2019) “Deep neural networks with transfer learning in millet crop images,” *Computers in Industry*, 108, pp. 115–120. Available at: <https://doi.org/10.1016/j.compind.2019.02.003>.

Crane-Droesch, A. (2018) “Machine learning methods for crop yield prediction and climate change impact assessment in agriculture,” *Environmental Research Letters*, 13(11), p. 114003. Available at: <https://doi.org/10.1088/1748-9326/aae159>.

Crevier, D. (1993) *AI: the tumultuous history of the search for artificial intelligence*. New York: Basic books.

Cruz, A.C. *et al.* (2017) “X-FIDO: An Effective Application for Detecting Olive Quick Decline Syndrome with Deep Learning and Data Fusion,” *Frontiers in Plant Science*, 8, p. 1741. Available at: <https://doi.org/10.3389/fpls.2017.01741>.

Ćuk, N. *et al.* (2022) “Introducing a cut-stem inoculation method for fast evaluation of sunflower resistance to *Macrophomina phaseolina*,” *Phytoparasitica*, 50(4), pp. 775–788. Available at: <https://doi.org/10.1007/s12600-022-01015-0>.

Cunningham, P., Cord, M. and Delany, S.J. (2008) “Supervised Learning,” in M. Cord and P. Cunningham (eds.) *Machine Learning Techniques for Multimedia*. Berlin, Heidelberg: Springer Berlin Heidelberg (Cognitive

Technologies), pp. 21–49. Available at: https://doi.org/10.1007/978-3-540-75171-7_2.

Curran, P.J. (1987) “Remote Sensing in Agriculture: An Introductory Review,” *Journal of Geography*, 86(4), pp. 147–156. Available at: <https://doi.org/10.1080/00221348708979166>.

Cvejić, S. *et al.* (2023) “Oil yield prediction for sunflower hybrid selection using different machine learning algorithms,” *Scientific Reports*, 13(1), p. 17611. Available at: <https://doi.org/10.1038/s41598-023-44999-3>.

Dartmann, G. (ed.) (2021) *Smart transportation: AI enabled mobility and autonomous driving*. First edition. Boca Raton, FL: CRC Press.

Das, S. *et al.* (2021) “Evaluation of water status of wheat genotypes to aid prediction of yield on sodic soils using UAV-thermal imaging and machine learning,” *Agricultural and Forest Meteorology*, 307, p. 108477. Available at: <https://doi.org/10.1016/j.agrformet.2021.108477>.

Dedić, B. *et al.* (2022) “Designing Sunflower for Biotic Stress Resilience: Everlasting Challenge,” in C. Kole (ed.) *Genomic Designing for Biotic Stress Resistant Oilseed Crops*. Cham: Springer International Publishing, pp. 85–136. Available at: https://doi.org/10.1007/978-3-030-91035-8_3.

Deng, J. *et al.* (2009) “ImageNet: A large-scale hierarchical image database,” *2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255. Available at: <https://doi.org/10.1109/CVPR.2009.5206848>.

Douarre, C. *et al.* (2018) “Transfer Learning from Synthetic Data Applied to Soil–Root Segmentation in X-Ray Tomography Images,” *Journal of Imaging*, 4(5), p. 65. Available at: <https://doi.org/10.3390/jimaging4050065>.

Elavarasan, D. *et al.* (2020) “A Hybrid CFS Filter and RF-RFE Wrapper-Based Feature Extraction for Enhanced Agricultural Crop Yield Prediction Modeling,” *Agriculture*, 10(9), p. 400. Available at: <https://doi.org/10.3390/agriculture10090400>.

Elavarasan, D. and Vincent, P.M.D. (2020) “Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications,” *IEEE Access*, 8, pp. 86886–86901. Available at: <https://doi.org/10.1109/ACCESS.2020.2992480>.

Esteva, A. *et al.* (2017) “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, 542(7639), pp. 115–118. Available at: <https://doi.org/10.1038/nature21056>.

Farah, A.A. *et al.* (2025) “The multifaceted impact of climate change on agricultural productivity: a systematic literature review of SCOPUS-indexed studies (2015–2024),” *Discover Sustainability*, 6(1), p. 397. Available at: <https://doi.org/10.1007/s43621-025-01229-2>.

Farooq, M.S. *et al.* (2020) “Role of IoT Technology in Agriculture: A Systematic Literature Review,” *Electronics*, 9(2), p. 319. Available at: <https://doi.org/10.3390/electronics9020319>.

Ferentinos, K.P. (2018) “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, 145, pp. 311–318. Available at: <https://doi.org/10.1016/j.compag.2018.01.009>.

Firat, M. (2023) “Integrating AI Applications into Learning Management Systems to Enhance e-Learning,” *Öğretim Teknolojisi ve Hayat Boyu Öğrenme Dergisi - Instructional Technology and Lifelong Learning*, pp. 1–14. Available at: <https://doi.org/10.52911/itall.1244453>.

Fuentes, A. *et al.* (2017) “A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition,” *Sensors*, 17(9), p. 2022. Available at: <https://doi.org/10.3390/s17092022>.

Fujita, E. *et al.* (2016) “Basic Investigation on a Robust and Practical Plant Diagnostic System,” *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Anaheim, CA, USA: IEEE, pp. 989–992. Available at: <https://doi.org/10.1109/ICMLA.2016.0178>.

Gatys, L.A., Ecker, A.S. and Bethge, M. (2015) “A Neural Algorithm of Artistic Style.” arXiv. Available at: <https://doi.org/10.48550/arXiv.1508.06576>.

Gebbers, R. and Adamchuk, V.I. (2010) “Precision Agriculture and Food Security,” *Science*, 327(5967), pp. 828–831. Available at: <https://doi.org/10.1126/science.1183899>.

Ghosh, P. *et al.* (2023) “Recognition of Sunflower Diseases Using Hybrid Deep Learning and Its Explainability with AI,” *Mathematics*, 11(10), p. 2241. Available at: <https://doi.org/10.3390/math11102241>.

Gleick, P.H. *et al.* (2014) *The World’s Water Volume 8: The Biennial Report on Freshwater Resources*. Island Press.

Goertzel, B. (2014) “Artificial General Intelligence: Concept, State of the Art, and Future Prospects,” *Journal of Artificial General Intelligence*, 5(1), pp. 1–48. Available at: <https://doi.org/10.2478/jagi-2014-0001>.

Goertzel, B. and Pennachin, C. (eds.) (2007) *Artificial General Intelligence*. Berlin, Heidelberg: Springer Berlin Heidelberg (Cognitive Technologies). Available at: <https://doi.org/10.1007/978-3-540-68677-4>.

Gonzalez-Sanchez, A., Frausto-Solis, J. and Ojeda-Bustamante, W. (2014) “Predictive ability of machine learning methods for massive crop yield prediction,” *Spanish Journal of Agricultural Research*, 12(2), pp. 313–328. Available at: <https://doi.org/10.5424/sjar/2014122-4439>.

Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep learning*. Cambridge, Mass: The MIT press (Adaptive computation and machine learning).

Goyal, P. *et al.* (2024) “A deep learning approach for early detection of drought stress in maize using proximal scale digital images,” *Neural Computing and Applications*, 36(4), pp. 1899–1913. Available at: <https://doi.org/10.1007/s00521-023-09219-z>.

Graves, A. (2014) “Generating Sequences With Recurrent Neural Networks.” arXiv. Available at: <https://doi.org/10.48550/arXiv.1308.0850>.

Gulzar, Y. *et al.* (2023) “Harnessing the Power of Transfer Learning in Sunflower Disease Detection: A Comparative Study,” *Agriculture*, 13(8), p. 1479. Available at: <https://doi.org/10.3390/agriculture13081479>.

Guo, W.W. and Xue, H. (2014) “Crop Yield Forecasting Using Artificial Neural Networks: A Comparison between Spatial and Temporal Models,” *Mathematical Problems in Engineering*. Edited by C.-C. Hung, 2014(1), p. 857865. Available at: <https://doi.org/10.1155/2014/857865>.

Guo, Y. *et al.* (2021) “Integrated phenology and climate in rice yields prediction using machine learning methods,” *Ecological Indicators*, 120, p. 106935. Available at: <https://doi.org/10.1016/j.ecolind.2020.106935>.

Gutierrez, J. *et al.* (2014) “Automated Irrigation System Using a Wireless Sensor Network and GPRS Module,” *IEEE Transactions on Instrumentation and Measurement*, 63(1), pp. 166–176. Available at: <https://doi.org/10.1109/TIM.2013.2276487>.

Hagen, M. *et al.* (2024) “Interpretable machine learning decodes soil microbiome’s response to drought stress,” *Environmental Microbiome*, 19(1), p. 35. Available at: <https://doi.org/10.1186/s40793-024-00578-1>.

Halevy, A., Norvig, P. and Pereira, F. (2009) “The Unreasonable Effectiveness of Data,” *IEEE Intelligent Systems*, 24(2), pp. 8–12. Available at: <https://doi.org/10.1109/MIS.2009.36>.

Halkiopoulos, C. and Gkintoni, E. (2024) “Leveraging AI in E-Learning: Personalized Learning and Adaptive Assessment through Cognitive Neuropsychology—A Systematic Analysis,” *Electronics*, 13(18), p. 3762. Available at: <https://doi.org/10.3390/electronics13183762>.

Hannun, A. *et al.* (2014) “Deep Speech: Scaling up end-to-end speech recognition.” arXiv. Available at: <https://doi.org/10.48550/arXiv.1412.5567>.

Haque, Md.A. and Li, S. (2025) “Exploring ChatGPT and its impact on society,” *AI and Ethics*, 5(2), pp. 791–803. Available at: <https://doi.org/10.1007/s43681-024-00435-4>.

Haralick, R.M., Shanmugam, K. and Dinstein, I. (1973) “Textural Features for Image Classification,” *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), pp. 610–621. Available at: <https://doi.org/10.1109/TSMC.1973.4309314>.

Hastie, T., Tibshirani, R. and Friedman, J.H. (2017) *The elements of statistical learning: data mining, inference, and prediction*. Second edition. New York, NY: Springer (Springer Series in Statistics).

He, K. *et al.* (2016) “Deep Residual Learning for Image Recognition.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778. Available at: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html (Accessed: June 22, 2026).

Hintze, A. (2016) *Understanding the four types of AI, from reactive robots to self-aware beings, The Conversation*. Available at: <https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616> (Accessed: November 30, 2025).

Huang, G. *et al.* (2017) “Densely Connected Convolutional Networks.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708. Available at: https://openaccess.thecvf.com/content_cvpr_2017/html/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.html (Accessed: June 22, 2026).

Hussain, M. *et al.* (2024) “AI-Optimized Irrigation for Sustainable Agriculture,” *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*. *2024 International Conference on Advances in Data Engineering and Intelligent Computing*

Systems (ADICS), Chennai, India: IEEE, pp. 01–09. Available at: <https://doi.org/10.1109/ADICS58448.2024.10533562>.

James, G. *et al.* (2017) *An introduction to statistical learning: with applications in R*. Corrected at 8th printing 2017. New York: Springer (Springer texts in statistics).

Jung, A. (2022) *Machine learning: the basics*. Singapore: Springer Singapore (Machine learning: foundations, methodologies, and applications).

Kamarudin, Mh. and Ismail, Z.H. (2022) “Lightweight deep CNN models for identifying drought stressed plant,” *IOP Conference Series: Earth and Environmental Science*, 1091(1), p. 012043. Available at: <https://doi.org/10.1088/1755-1315/1091/1/012043>.

Kamarudin, M.H., Ismail, Z.H. and Saidi, N.B. (2021) “Deep Learning Sensor Fusion in Plant Water Stress Assessment: A Comprehensive Review,” *Applied Sciences*, 11(4), p. 1403. Available at: <https://doi.org/10.3390/app11041403>.

Kanellopoulos, I. and Wilkinson, G.G. (1997) “Strategies and best practice for neural network image classification,” *International Journal of Remote Sensing*, 18(4), pp. 711–725. Available at: <https://doi.org/10.1080/014311697218719>.

Karim, Foughali, Karim, Fathalah and Frihida, A. (2017) “Monitoring system using web of things in precision agriculture,” *Procedia Computer Science*, 110, pp. 402–409. Available at: <https://doi.org/10.1016/j.procs.2017.06.083>.

Karpathy, A. and Fei-Fei, L. (2015) “Deep Visual-Semantic Alignments for Generating Image Descriptions.” Boston, MA, USA: IEEE, pp. 3128–3137. Available at: https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Karpathy_Deep_Visual-Semantic_Alignments_2015_CVPR_paper.html.

Keras: Deep Learning for humans (no date). Available at: <https://keras.io/> (Accessed: April 17, 2026).

Khaki, S. and Wang, L. (2019) “Crop Yield Prediction Using Deep Neural Networks,” *Frontiers in Plant Science*, 10, p. 621. Available at: <https://doi.org/10.3389/fpls.2019.00621>.

Khaki, S., Wang, L. and Archontoulis, S.V. (2020) “A CNN-RNN Framework for Crop Yield Prediction,” *Frontiers in Plant Science*, 10, p. 1750. Available at: <https://doi.org/10.3389/fpls.2019.01750>.

Khalifani, S. *et al.* (2022) “Prediction of sunflower grain yield under normal and salinity stress by RBF, MLP and, CNN models,” *Industrial Crops and Products*, 189, p. 115762. Available at: <https://doi.org/10.1016/j.indcrop.2022.115762>.

Kikon, A. and Deka, P.C. (2022) “Artificial intelligence application in drought assessment, monitoring and forecasting: a review,” *Stochastic Environmental Research and Risk Assessment*, 36(5), pp. 1197–1214. Available at: <https://doi.org/10.1007/s00477-021-02129-3>.

Kingma, D.P. and Ba, J. (2017) “Adam: A Method for Stochastic Optimization.” arXiv. Available at: <https://doi.org/10.48550/arXiv.1412.6980>.

Klepper, B. and Kaspar, T.C. (1994) “Rhizotrons: Their Development and Use in Agricultural Research,” *Agronomy Journal*, 86(5), pp. 745–753. Available at: <https://doi.org/10.2134/agronj1994.00021962008600050002x>.

Kober, J., Bagnell, J.A. and Peters, J. (2013) “Reinforcement learning in robotics: A survey,” *The International Journal of Robotics Research*, 32(11), pp. 1238–1274. Available at: <https://doi.org/10.1177/0278364913495721>.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) “ImageNet Classification with Deep Convolutional Neural Networks,” *Advances in Neural Information Processing Systems*. Curran Associates, Inc. Available at: https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html (Accessed: June 22, 2026).

Kuo, C.-E. *et al.* (2023) “Early detection of drought stress in tomato from spectroscopic data: A novel convolutional neural network with feature selection,” *Chemometrics and Intelligent Laboratory Systems*, 239, p. 104869. Available at: <https://doi.org/10.1016/j.chemolab.2023.104869>.

Kurtulmuş, F. (2021) “Identification of sunflower seeds with deep convolutional neural networks,” *Journal of Food Measurement and Characterization*, 15(2), pp. 1024–1033. Available at: <https://doi.org/10.1007/s11694-020-00707-7>.

Lazić, O. *et al.* (2024) “Transfer Learning in Multimodal Sunflower Drought Stress Detection,” *Applied Sciences*, 14(14), p. 6034. Available at: <https://doi.org/10.3390/app14146034>.

LeCun, Y. *et al.* (1990) “Handwritten digit recognition with a back-propagation network.” *Advances in Neural Information Processing Systems*.

- Lecun, Y. *et al.* (1998) “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, 86(11), pp. 2278–2324. Available at: <https://doi.org/10.1109/5.726791>.
- LeCun, Y., Bengio, Y. and Hinton, G. (2015) “Deep learning,” *Nature*, 521(7553), pp. 436–444. Available at: <https://doi.org/10.1038/nature14539>.
- Lee, I. and Shin, Y.J. (2020) “Machine learning for enterprises: Applications, algorithm selection, and challenges,” *Business Horizons*, 63(2), pp. 157–170. Available at: <https://doi.org/10.1016/j.bushor.2019.10.005>.
- Lee, J. (2020) *Industrial AI: Applications with Sustainable Performance*. Singapore: Springer Singapore. Available at: <https://doi.org/10.1007/978-981-15-2144-7>.
- Leminen Madsen, S. *et al.* (2020) “Open Plant Phenotype Database of Common Weeds in Denmark,” *Remote Sensing*, 12(8), p. 1246. Available at: <https://doi.org/10.3390/rs12081246>.
- Li, W. *et al.* (2022) “A new lightweight deep neural network for surface scratch detection,” *The International Journal of Advanced Manufacturing Technology*, 123(5–6), pp. 1999–2015. Available at: <https://doi.org/10.1007/s00170-022-10335-8>.
- Li, Y. *et al.* (2009) “Climate change and drought: a risk assessment of crop-yield impacts,” *Climate Research*, 39, pp. 31–46. Available at: <https://doi.org/10.3354/cr00797>.
- Lin, D., Wei, R. and Xu, L. (2019) “An Integrated Yield Prediction Model for Greenhouse Tomato,” *Agronomy*, 9(12), p. 873. Available at: <https://doi.org/10.3390/agronomy9120873>.
- Lindblom, J. *et al.* (2017) “Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies,” *Precision Agriculture*, 18(3), pp. 309–331. Available at: <https://doi.org/10.1007/s11119-016-9491-4>.
- Liu, B. *et al.* (2017) “Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks,” *Symmetry*, 10(1), p. 11. Available at: <https://doi.org/10.3390/sym10010011>.
- Liu, X. *et al.* (2019) “A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis,” *The Lancet Digital Health*, 1(6), pp. e271–e297. Available at: [https://doi.org/10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2).

Lu, J. *et al.* (2017) “An in-field automatic wheat disease diagnosis system,” *Computers and Electronics in Agriculture*, 142, pp. 369–379. Available at: <https://doi.org/10.1016/j.compag.2017.09.012>.

Lužaić, T. *et al.* (2023) “Investigation of oxidative characteristics, fatty acid composition and bioactive compounds content in cold pressed oils of sunflower grown in Serbia and Argentina,” *Heliyon*, 9(7), p. e18201. Available at: <https://doi.org/10.1016/j.heliyon.2023.e18201>.

Ma, Y. *et al.* (2020) “Artificial intelligence applications in the development of autonomous vehicles: a survey,” *IEEE/CAA Journal of Automatica Sinica*, 7(2), pp. 315–329. Available at: <https://doi.org/10.1109/JAS.2020.1003021>.

MacQueen, J. (1967) “Some methods for classification and analysis of multivariate observations,” *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. University of California Press, pp. 281–298. Available at: <https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fifth-Berkeley-Symposium-on-Mathematical-Statistics-and/chapter/Some-methods-for-classification-and-analysis-of-multivariate-observations/bsmsp/1200512992> (Accessed: June 24, 2026).

Maffezzoli, F. *et al.* (2022) “Agriculture 4.0: A systematic literature review on the paradigm, technologies and benefits,” *Futures*, 142, p. 102998. Available at: <https://doi.org/10.1016/j.futures.2022.102998>.

Mahat, D. *et al.* (2023) “AI-Driven Optimization of Supply Chain and Logistics in Mechanical Engineering,” *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*. 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Gautam Buddha Nagar, India: IEEE, pp. 1611–1616. Available at: <https://doi.org/10.1109/UPCON59197.2023.10434905>.

Malik, A. *et al.* (2022) “Design and Evaluation of a Hybrid Technique for Detecting Sunflower Leaf Disease Using Deep Learning Approach,” *Journal of Food Quality*. Edited by M. Faisal Manzoor, 2022, pp. 1–12. Available at: <https://doi.org/10.1155/2022/9211700>.

Maya Gopal, P.S. and Bhargavi, R. (2019) “A novel approach for efficient crop yield prediction,” *Computers and Electronics in Agriculture*, 165, p. 104968. Available at: <https://doi.org/10.1016/j.compag.2019.104968>.

Mbewu, A.D. *et al.* (2024) “Climate change effects on water footprint of crop production: A meta-analysis,” *Environmental Challenges*, 17, p. 101033. Available at: <https://doi.org/10.1016/j.envc.2024.101033>.

McCarthy, J. *et al.* (1955) *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Dartmouth College.

McCulloch, W.S. and Pitts, W. (1943) “A logical calculus of the ideas immanent in nervous activity,” *The Bulletin of Mathematical Biophysics*, 5(4), pp. 115–133. Available at: <https://doi.org/10.1007/BF02478259>.

McKinney, S.M. *et al.* (2020) “International evaluation of an AI system for breast cancer screening,” *Nature*, 577(7788), pp. 89–94. Available at: <https://doi.org/10.1038/s41586-019-1799-6>.

Mittelbach, H. *et al.* (2011) “Soil moisture monitoring for climate research: Evaluation of a low-cost sensor in the framework of the Swiss Soil Moisture Experiment (SwissSMEX) campaign,” *Journal of Geophysical Research*, 116(D5), p. D05111. Available at: <https://doi.org/10.1029/2010JD014907>.

Mnih, V. *et al.* (2015) “Human-level control through deep reinforcement learning,” *Nature*, 518(7540), pp. 529–533. Available at: <https://doi.org/10.1038/nature14236>.

Mohanty, S.P., Hughes, D.P. and Salathé, M. (2016) “Using Deep Learning for Image-Based Plant Disease Detection,” *Frontiers in Plant Science*, 7, p. 1419. Available at: <https://doi.org/10.3389/fpls.2016.01419>.

Mordvintsev, A., Olah, C. and Tyka, M. (2015) *Inceptionism: Going deeper into neural networks*, Google Research Blog. Available at: <https://research.google/blog/inceptionism-going-deeper-into-neural-networks/> (Accessed: December 3, 2025).

Mukherjee, S., Mishra, A. and Trenberth, K.E. (2018) “Climate Change and Drought: a Perspective on Drought Indices,” *Current Climate Change Reports*, 4(2), pp. 145–163. Available at: <https://doi.org/10.1007/s40641-018-0098-x>.

Mulla, D.J. (2013) “Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps,” *Biosystems Engineering*, 114(4), pp. 358–371. Available at: <https://doi.org/10.1016/j.biosystemseng.2012.08.009>.

Muñoz, A. (2014) *Machine Learning and Optimization*. Course Notes / Technical Report. New York, USA (opciono): Courant Institute of Mathematical Sciences, pp. 1–2.

Murphy, K.P. (2013) *Machine learning: a probabilistic perspective*. 4. print. (fixed many typos). Cambridge, Mass.: MIT Press (Adaptive computation and machine learning series).

Nagel, K.A. *et al.* (2012) “GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons,” *Functional Plant Biology*, 39(11), pp. 891–904. Available at: <https://doi.org/10.1071/FP12023>.

Nagib, A.E. *et al.* (2022) “Hyperparameters Optimization of Deep Convolutional Neural Network for Detecting COVID-19 Using Differential Evolution,” in S.A. Hassan, A.W. Mohamed, and K.A. Alnowibet (eds.) *Decision Sciences for COVID-19: Learning Through Case Studies*. Cham: Springer International Publishing, pp. 305–325. Available at: https://doi.org/10.1007/978-3-030-87019-5_18.

Navarro-Hellín, H. *et al.* (2016) “A decision support system for managing irrigation in agriculture,” *Computers and Electronics in Agriculture*, 124, pp. 121–131. Available at: <https://doi.org/10.1016/j.compag.2016.04.003>.

Nazir, A. and Wang, Z. (2023) “A comprehensive survey of ChatGPT: Advancements, applications, prospects, and challenges,” *Meta-Radiology*, 1(2), p. 100022. Available at: <https://doi.org/10.1016/j.metrad.2023.100022>.

Newborn, M. (2003) *Deep Blue*. New York, NY: Springer New York. Available at: <https://doi.org/10.1007/978-0-387-21790-1>.

Newell, A., Shaw, J.C. and Simon, H.A. (1957) “Empirical explorations of the logic theory machine: a case study in heuristic,” *Papers presented at the February 26-28, 1957, western joint computer conference: Techniques for reliability*. New York, NY, USA: Association for Computing Machinery (IRE-AIEE-ACM '57 (Western)), pp. 218–230. Available at: <https://doi.org/10.1145/1455567.1455605>.

Oliveira, R.C.D. and Silva, R.D.D.S.E. (2023) “Artificial Intelligence in Agriculture: Benefits, Challenges, and Trends,” *Applied Sciences*, 13(13), p. 7405. Available at: <https://doi.org/10.3390/app13137405>.

P. S., M.G. and R., B. (2019) “Performance Evaluation of Best Feature Subsets for Crop Yield Prediction Using Machine Learning Algorithms,” *Applied Artificial Intelligence*, 33(7), pp. 621–642. Available at: <https://doi.org/10.1080/08839514.2019.1592343>.

Pan, S.J. and Yang, Q. (2010) “A Survey on Transfer Learning,” *IEEE Transactions on Knowledge and Data Engineering*, 22(10), pp. 1345–1359. Available at: <https://doi.org/10.1109/TKDE.2009.191>.

Panda, S.S., Ames, D.P. and Panigrahi, S. (2010) “Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques,” *Remote Sensing*, 2(3), pp. 673–696. Available at: <https://doi.org/10.3390/rs2030673>.

Patra, A.K. and Sahoo, L. (2024) “Explainable light-weight deep learning pipeline for improved drought stress identification,” *Frontiers in Plant Science*, 15, p. 1476130. Available at: <https://doi.org/10.3389/fpls.2024.1476130>.

Peng, B. *et al.* (2020) “Assessing the benefit of satellite-based Solar-Induced Chlorophyll Fluorescence in crop yield prediction,” *International Journal of Applied Earth Observation and Geoinformation*, 90, p. 102126. Available at: <https://doi.org/10.1016/j.jag.2020.102126>.

Penmetsa, A.V. and Sarma, T.H. (2022) “Crop Type and Stress Detection using Transfer Learning with MobileNet,” *2022 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)*. *2022 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)*, Hyderabad, India: IEEE, pp. 1–4. Available at: <https://doi.org/10.1109/ICETCI55171.2022.9921356>.

Pettorelli, N. *et al.* (2005) “Using the satellite-derived NDVI to assess ecological responses to environmental change,” *Trends in Ecology & Evolution*, 20(9), pp. 503–510. Available at: <https://doi.org/10.1016/j.tree.2005.05.011>.

Pound, M.P. *et al.* (2017) “Deep machine learning provides state-of-the-art performance in image-based plant phenotyping,” *GigaScience*, 6(10), pp. 1–10. Available at: <https://doi.org/10.1093/gigascience/gix083>.

Ramcharan, A. *et al.* (2017) “Deep Learning for Image-Based Cassava Disease Detection,” *Frontiers in Plant Science*, 8, p. 1852. Available at: <https://doi.org/10.3389/fpls.2017.01852>.

Rao, P.V., Gayathri, R. and Sunitha, R. (2015) “A Novel Approach for Design and Analysis of Diabetic Retinopathy Glaucoma Detection Using Cup to Disk Ratio and ANN,” *Procedia Materials Science*, 10, pp. 446–454. Available at: <https://doi.org/10.1016/j.mspro.2015.06.080>.

Rojanarungruengporn, K. and Pumrin, S. (2021) “Early Stress Detection in Plant Phenotyping using CNN and LSTM Architecture,” *2021 9th International Electrical Engineering Congress (iEECON)*. *2021 9th International Electrical Engineering Congress (iEECON)*, Pattaya, Thailand: IEEE, pp. 389–392. Available at: <https://doi.org/10.1109/iEECON51072.2021.9440342>.

Rose, D.C. *et al.* (2016) “Decision support tools for agriculture: Towards effective design and delivery,” *Agricultural Systems*, 149, pp. 165–174. Available at: <https://doi.org/10.1016/j.agsy.2016.09.009>.

Rosenblatt, F. (1958) “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review*, 65(6), pp. 386–408. Available at: <https://doi.org/10.1037/h0042519>.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) “Learning representations by back-propagating errors,” *Nature*, 323(6088), pp. 533–536. Available at: <https://doi.org/10.1038/323533a0>.

Russell, S.J. and Norvig, P. (2016) *Artificial intelligence: a modern approach*. Third edition, Global edition. Boston Columbus Indianapolis: Pearson (Prentice Hall series in artificial intelligence).

Sabo, F., Meroni, M., Waldner, F., & Rembold, F. (2023). Is deeper always better? Evaluating deep learning models for yield forecasting with small data. *Environmental Monitoring and Assessment*, 195(10), 1153.

Sahoo, K., Samal, A. K., Pramanik, J. and Pani, S. K. (2019) “Exploratory data analysis using Python,” *International Journal of Innovative Technology and Exploring Engineering*, 8(12), pp. 4727–4735.

Sarker, I.H. (2021) “Machine Learning: Algorithms, Real-World Applications and Research Directions,” *SN Computer Science*, 2(3), p. 160. Available at: <https://doi.org/10.1007/s42979-021-00592-x>.

Satir, O. and Berberoglu, S. (2016) “Crop yield prediction under soil salinity using satellite derived vegetation indices,” *Field Crops Research*, 192, pp. 134–143. Available at: <https://doi.org/10.1016/j.fcr.2016.04.028>.

Setiowati, S. *et al.* (2017) “A review of optimization method in face recognition: Comparison deep learning and non-deep learning methods,” *2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE)*. *2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE)*, Phuket: IEEE, pp. 1–6. Available at: <https://doi.org/10.1109/ICITEED.2017.8250484>.

Shahhosseini, M., Hu, G. and Archontoulis, S.V. (2020) “Forecasting Corn Yield With Machine Learning Ensembles,” *Frontiers in Plant Science*, 11, p. 1120. Available at: <https://doi.org/10.3389/fpls.2020.01120>.

Sharma, H., Sidhu, H. and Bhowmik, A. (2025) “Remote Sensing Using Unmanned Aerial Vehicles for Water Stress Detection: A Review Focusing

on Specialty Crops,” *Drones*, 9(4), p. 241. Available at: <https://doi.org/10.3390/drones9040241>.

Shen, R. *et al.* (2019) “Construction of a drought monitoring model using deep learning based on multi-source remote sensing data,” *International Journal of Applied Earth Observation and Geoinformation*, 79, pp. 48–57. Available at: <https://doi.org/10.1016/j.jag.2019.03.006>.

Shin, H.-C. *et al.* (2016) “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,” *IEEE transactions on medical imaging*, 35(5), pp. 1285–1298. Available at: <https://doi.org/10.1109/TMI.2016.2528162>.

Shorten, C., Khoshgoftaar, T.M. and Furht, B. (2021) “Text Data Augmentation for Deep Learning,” *Journal of Big Data*, 8(1), p. 101. Available at: <https://doi.org/10.1186/s40537-021-00492-0>.

Siciliano, B. and Khatib, O. (2016) “Robotics and the Handbook,” in B. Siciliano and O. Khatib (eds.) *Springer Handbook of Robotics*. Cham: Springer International Publishing (Springer Handbooks), pp. 1–6. Available at: https://doi.org/10.1007/978-3-319-32552-1_1.

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., *et al.* (2016) “Mastering the game of Go with deep neural networks and tree search,” *Nature*, 529(7587), pp. 484–489. Available at: <https://doi.org/10.1038/nature16961>.

Simonyan, K. and Zisserman, A. (2015) “Very Deep Convolutional Networks for Large-Scale Image Recognition.” arXiv. Available at: <https://doi.org/10.48550/arXiv.1409.1556>.

Singh, A.K. *et al.* (2018) “Deep Learning for Plant Stress Phenotyping: Trends and Future Perspectives,” *Trends in Plant Science*, 23(10), pp. 883–898. Available at: <https://doi.org/10.1016/j.tplants.2018.07.004>.

Singh, R. (2024) *Convolutional Neural Networks (CNNs) in a minute, Medium*. Available at: <https://ai.plainenglish.io/convolutional-neural-networks-cnns-in-a-minute-4c9ae1b9a2cf> (Accessed: June 21, 2026).

Sladojevic, S. *et al.* (2016) “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,” *Computational Intelligence and Neuroscience*, 2016, pp. 1–11. Available at: <https://doi.org/10.1155/2016/3289801>.

Suharso, A. *et al.* (2022) “The Role of Machine Learning in Remote Sensing for Agriculture Drought Monitoring: A Systematic Review,” *International*

Journal of Advanced Computer Science and Applications (IJACSA), 13(12). Available at: <https://doi.org/10.14569/IJACSA.2022.0131290>.

Sun, J. *et al.* (2019) “County-Level Soybean Yield Prediction Using Deep CNN-LSTM Model,” *Sensors*, 19(20), p. 4363. Available at: <https://doi.org/10.3390/s19204363>.

Sutton, R.S. and Barto, A.G. (2018) Reinforcement Learning: An Introduction. 2nd Edition, A Bradford Book, Cambridge, MA.

Strauss, S.Y. *et al.* (2002) “Direct and ecological costs of resistance to herbivory,” *Trends in Ecology & Evolution*, 17(6), pp. 278–285. Available at: [https://doi.org/10.1016/S0169-5347\(02\)02483-7](https://doi.org/10.1016/S0169-5347(02)02483-7).

Swaminathan, B. and Vairavasundaram, S. (2024) “D2CNN: Double-staged deep CNN for stress identification and classification in cropping system,” *Agricultural Systems*, 216, p. 103886. Available at: <https://doi.org/10.1016/j.agsy.2024.103886>.

Szegedy, C. *et al.* (2015) “Going Deeper With Convolutions.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9. Available at: https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Szegedy_Going_Deep_With_2015_CVPR_paper.html (Accessed: June 21, 2026).

Škorić, D. & Jocić, S. Achievements of sunflower breeding at the IFVC in Novi Sad. In Proc. 16th Intl. Sunflower Conf., Fargo, ND, USA, 29, pp. 441–448 (2004).

Talaviya, T. *et al.* (2020) “Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides,” *Artificial Intelligence in Agriculture*, 4, pp. 58–73. Available at: <https://doi.org/10.1016/j.aiia.2020.04.002>.

Talukdar, J., Biswas, A. and Gupta, S. (2018) “Data Augmentation on Synthetic Images for Transfer Learning using Deep CNNs,” *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*. 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN), Noida: IEEE, pp. 215–219. Available at: <https://doi.org/10.1109/SPIN.2018.8474209>.

Tan, M. and Le, Q.V. (2020) “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.” arXiv. Available at: <https://doi.org/10.48550/arXiv.1905.11946>.

Tang, Z. *et al.* (2020) “OnlineAugment: Online Data Augmentation with Less Domain Knowledge,” in A. Vedaldi *et al.* (eds.) *Computer Vision – ECCV 2020*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 313–329. Available at: https://doi.org/10.1007/978-3-030-58571-6_19.

Tejasri, N. *et al.* (2023) “Intelligent Drought Stress Monitoring on Spatio- Spectral- Temporal Drone based Crop Imagery using Deep Networks.” *2nd AAAI Workshop on AI for Agriculture and Food Systems*. Available at: <https://openreview.net/pdf?id=6U08SGo64>.

TensorFlow Developers (2022). TensorFlow Hub. [Online]. Available: <https://tfhub.dev>

Thapa, R. *et al.* (2020) “The Plant Pathology Challenge 2020 data set to classify foliar disease of apples,” *Applications in Plant Sciences*, 8(9), p. e11390. Available at: <https://doi.org/10.1002/aps3.11390>.

Tian, D. *et al.* (2003) “Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*,” *Nature*, 423(6935), pp. 74–77. Available at: <https://doi.org/10.1038/nature01588>.

Tizhoosh, H.R. and Pantanowitz, L. (2018) “Artificial Intelligence and Digital Pathology: Challenges and Opportunities,” *Journal of Pathology Informatics*, 9(1), p. 38. Available at: https://doi.org/10.4103/jpi.jpi_53_18.

Tsang, Y.P. and Lee, C.K.M. (2022) “Artificial intelligence in industrial design: A semi-automated literature survey,” *Engineering Applications of Artificial Intelligence*, 112, p. 104884. Available at: <https://doi.org/10.1016/j.engappai.2022.104884>.

Tukey, J. W. (1977) *Exploratory Data Analysis*. Reading, Massachusetts: Addison-Wesley.

TURING, A.M. (1950) “I.—COMPUTING MACHINERY AND INTELLIGENCE,” *Mind*, LIX(236), pp. 433–460. Available at: <https://doi.org/10.1093/mind/LIX.236.433>.

Vear, F. (2016) “Changes in sunflower breeding over the last fifty years,” *OCL*, 23(2), p. D202. Available at: <https://doi.org/10.1051/ocl/2016006>.

Viani, F. *et al.* (2017) “Low-Cost Wireless Monitoring and Decision Support for Water Saving in Agriculture,” *IEEE Sensors Journal*, 17(13), pp. 4299–4309. Available at: <https://doi.org/10.1109/JSEN.2017.2705043>.

Virnodkar, S.S. *et al.* (2020) “Remote sensing and machine learning for crop water stress determination in various crops: a critical review,” *Precision Agriculture*, 21(5), pp. 1121–1155. Available at: <https://doi.org/10.1007/s11119-020-09711-9>.

Wang, Y. *et al.* (2020) “Combining Multi-Source Data and Machine Learning Approaches to Predict Winter Wheat Yield in the Conterminous United States,” *Remote Sensing*, 12(8), p. 1232. Available at: <https://doi.org/10.3390/rs12081232>.

Wang, Y. *et al.* (2023) “A Survey on ChatGPT: AI-Generated Contents, Challenges, and Solutions,” *IEEE Open Journal of the Computer Society*, 4, pp. 280–302. Available at: <https://doi.org/10.1109/OJCS.2023.3300321>.

Watkins, C.J.C.H. and Dayan, P. (1992) “Q-learning,” *Machine Learning*, 8(3), pp. 279–292. Available at: <https://doi.org/10.1007/BF00992698>.

Weizenbaum, J. (1966) “ELIZA—a computer program for the study of natural language communication between man and machine,” *Communications of the ACM*, 9(1), pp. 36–45. Available at: <https://doi.org/10.1145/365153.365168>.

Williamson, H.F. *et al.* (2023) “Data management challenges for artificial intelligence in plant and agricultural research,” *F1000Research*, 10, p. 324. Available at: <https://doi.org/10.12688/f1000research.52204.2>.

Wolfert, S. *et al.* (2017) “Big Data in Smart Farming – A review,” *Agricultural Systems*, 153, pp. 69–80. Available at: <https://doi.org/10.1016/j.agsy.2017.01.023>.

Xue, J. and Su, B. (2017) “Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications,” *Journal of Sensors*, 2017, pp. 1–17. Available at: <https://doi.org/10.1155/2017/1353691>.

Yalcin, H. (2019) “An Approximation for A Relative Crop Yield Estimate from Field Images Using Deep Learning,” *2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*. 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Istanbul, Turkey: IEEE, pp. 1–6. Available at: <https://doi.org/10.1109/Agro-Geoinformatics.2019.8820693>.

Yang, S. *et al.* (2022) “Transfer Learning from Synthetic In-vitro Soybean Pods Dataset for In-situ Segmentation of On-branch Soybean Pods,” *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2022 IEEE/CVF Conference on Computer Vision and

Pattern Recognition Workshops (CVPRW), New Orleans, LA, USA: IEEE, pp. 1665–1674. Available at: <https://doi.org/10.1109/CVPRW56347.2022.00173>.

Yosinski, J. *et al.* (2014) “How transferable are features in deep neural networks?” arXiv. Available at: <https://doi.org/10.48550/arXiv.1411.1792>.

Yuan, Y. *et al.* (2022) “Advanced agricultural disease image recognition technologies: A review,” *Information Processing in Agriculture*, 9(1), pp. 48–59. Available at: <https://doi.org/10.1016/j.inpa.2021.01.003>.

Yunseop Kim, Evans, R.G. and Iversen, W.M. (2008) “Remote Sensing and Control of an Irrigation System Using a Distributed Wireless Sensor Network,” *IEEE Transactions on Instrumentation and Measurement*, 57(7), pp. 1379–1387. Available at: <https://doi.org/10.1109/TIM.2008.917198>.

Zhai, X. *et al.* (2021) “A Review of Artificial Intelligence (AI) in Education from 2010 to 2020,” *Complexity*. Edited by N. Cai, 2021(1), p. 8812542. Available at: <https://doi.org/10.1155/2021/8812542>.

Zhang, L. *et al.* (2020) “Combining Optical, Fluorescence, Thermal Satellite, and Environmental Data to Predict County-Level Maize Yield in China Using Machine Learning Approaches,” *Remote Sensing*, 12(1), p. 21. Available at: <https://doi.org/10.3390/rs12010021>.

Zhang, N., Wang, M. and Wang, N. (2002) “Precision agriculture—a worldwide overview,” *Computers and Electronics in Agriculture*, 36(2), pp. 113–132. Available at: [https://doi.org/10.1016/S0168-1699\(02\)00096-0](https://doi.org/10.1016/S0168-1699(02)00096-0).

Zhang, W. *et al.* (1988) “Shift-Invariant Pattern Recognition Neural Network and Its Optical Architecture.” *Proceedings of Annual Conference of the Japan Society of Applied Physics*, Japan.

Zhang, W. *et al.* (1990) “Parallel distributed processing model with local space-invariant interconnections and its optical architecture,” *Applied Optics*, 29(32), p. 4790. Available at: <https://doi.org/10.1364/AO.29.004790>.

Zhao, J., Liu, D. and Huang, R. (2023) “A Review of Climate-Smart Agriculture: Recent Advancements, Challenges, and Future Directions,” *Sustainability*, 15(4), p. 3404. Available at: <https://doi.org/10.3390/su15043404>.

Zhou, Z.-H. (2012) *Ensemble methods: foundations and algorithms*. Online-Ausg. Boca Raton, FL: Taylor & Francis (Chapman & Hall).

Zhu, X. and Goldberg, A.B. (2009) “Overview of Semi-Supervised Learning,” in X. Zhu and A.B. Goldberg (eds.) *Introduction to Semi-Supervised Learning*.

Cham: Springer International Publishing, pp. 9–19. Available at:
https://doi.org/10.1007/978-3-031-01548-9_2.

Овај Образац чини саставни део докторске дисертације, односно докторског уметничког пројекта који се брани на Универзитету у Новом Саду. Попуњен Образац укоричити иза текста докторске дисертације, односно докторског уметничког пројекта.

План третмана података

Назив пројекта/истраживања
Примена вештачке интелигенције у оптимизацији узгоја сунцокрета: Детекција стреса услед суше и предвиђање висине приноса уља
Назив институције/институција у оквиру којих се спроводи истраживање
а) Факултет техничких наука, Универзитет у Новом Саду б) Институт за ратарство и повртарство, Нови Сад
Назив програма у оквиру ког се реализује истраживање
Докторске академске студије на Факултету техничких наука, Универзитет у Новом Саду.
1. Опис података
1.1 Врста студије <i>Укратко описати тип студије у оквиру које се подаци прикупљају</i> Истраживање обухвата експерименталну студију у области примене метода вештачке интелигенције у пољопривреди. Подаци су прикупљени у оквиру контролисаног експеримента за испитивање одговора сунцокрета на различите нивое доступности воде, као и у оквиру производних огледа сунцокрета, и коришћени су за развој и евалуацију модела за детекцију стреса услед суше и предвиђање висине приноса уља.
1.2 Врсте података а) квантитативни б) квалитативни
1.3. Начин прикупљања података а) анкете, упитници, тестови б) клиничке процене, медицински записи, електронски здравствени записи в) генотипови: генотипови/хибриди сунцокрета г) административни подаци: навести врсту _____

д) узорци ткива: навести врсту _____

ђ) снимци, фотографије: фотографије кореновог система и надземног дела биљака сунцокрета при различитим нивоима доступности воде

е) текст, навести врсту _____

ж) мапа, навести врсту _____

з) остало: агрономски и метеоролошки подаци прикупљени у оквиру производних огледа сунцокрета

1.3 Формат података, употребљене скале, количина података

1.3.1 Употребљени софтвер и формат датотеке:

а) Excel фајл, датотека .xlsx

б) SPSS фајл, датотека _____

в) PDF фајл, датотека _____

г) Текст фајл, датотека _____

е) JPG фајл, датотека .jpeg

ф) Остало, датотека _____

1.3.2. Број записа (код квантитативних података)

а) број варијабли: **8**

б) број мерења (испитаника, процена, снимака и сл.): **584 снимака; 29.699 квантитативних записа**

1.3.3. Поновљена мерења

а) да

б) не

Уколико је одговор да, одговорити на следећа питања:

а) временски размак измедју поновљених мера је _____

б) варијабле које се више пута мере односе се на _____

в) нове верзије фајлова који садрже поновљена мерења су именоване као _____

Напомене: _____

Да ли формати и софтвер омогућавају дељење и дугорочну валидност података?

а) Да

б) Не

Ако је одговор не, образложити _____

2. Прикупљање података

2.1 Методологија за прикупљање/генерисање података

2.1.1. У оквиру ког истраживачког нацрта су подаци прикупљени?

а) експеримент, навести тип: контролисани експеримент са различитим нивоима доступности воде и производни огледи сунцокрета на више локалитета.

б) корелационо истраживање, навести тип _____

ц) анализа текста, навести тип _____

д) остало, навести шта _____

2.1.2 Навести врсте мерних инструмената или стандарде података специфичних за одређену научну дисциплину (ако постоје).

За прикупљање података коришћени су Canon EOS 850D фотоапарат за снимање биљака, WinRhizo систем/софтвер за анализу кореновог система, Harvest Master HM800 и Wintersteiger Delta за мерење приноса, као и Maran Ultra-10 NMR анализатор за одређивање садржаја уља. За метеоролошке податке коришћени су подаци Републичког хидрометеоролошког завода Србије, релевантни за локалитете на којима су спроведени производни огледи сунцокрета.

2.2 Квалитет података и стандарди

2.2.1. Третман недостајућих података

а) Да ли матрица садржи недостајуће податке? Да Не

Ако је одговор да, одговорити на следећа питања:

а) Колики је број недостајућих података? 301

б) Да ли се кориснику матрице препоручује замена недостајућих података? Да Не

в) Ако је одговор да, навести сугестије за третман замене недостајућих података

2.2.2. На који начин је контролисан квалитет података? Описати

Квалитет података контролисан је провером потпуности и конзистентности података, идентификацијом недостајућих вредности и провером опсега и логичке исправности измерених вредности. За податке добијене снимањем биљака извршена је контрола квалитета и употребљивости снимака пре њиховог укључивања у анализу.

2.2.3. На који начин је извршена контрола уноса података у матрицу?

Контрола уноса података извршена је провером типова и опсега вредности, контролом конзистентности ознака генотипова/хибрида и експерименталних третмана, као и провером недостајућих и нелогичних вредности пре статистичке обраде и моделовања.

3. Третман података и пратећа документација

3.1. Третман и чување података

3.1.1. Подаци ће бити депоновани у _____ репозиторијум.

3.1.2. URL адреса _____

3.1.3. DOI _____

3.1.4. Да ли ће подаци бити у отвореном приступу?

а) Да

б) Да, али после ембарга који ће трајати до _____

в) Не

Ако је одговор не, навести разлог

Подаци представљају део истраживања реализованог у оквиру пројекта SMARTSUN, финансираног у оквиру програма ИДЕЈЕ Фонда за науку Републике Србије. Подаци су прикупљени и креирани у Институту за ратарство и повртарство у Новом Саду и остају у оквиру истраживачке инфраструктуре и система за управљање подацима Института. Из тог разлога, подаци нису предвиђени за јавно стављање у отворени приступ.

3.1.5. Подаци неће бити депоновани у репозиторијум, али ће бити чувани.

Образложење

Приступ подацима може бити омогућен истраживачима за потребе научноистраживачког рада, уз претходни договор са аутором података и институцијом која је податке ставила на располагање, у складу са условима под којима су подаци добијени и коришћени.

3.2 Метаподаци и документација података

3.2.1. Који стандард за метаподатке ће бити примењен? _____

3.2.1. Навести метаподатке на основу којих су подаци депоновани у репозиторијум.

Ако је потребно, навести методе које се користе за преузимање података, аналитичке и процедуралне информације, њихово кодирање, детаљне описе варијабли, записа итд.

3.3 Стратегија и стандарди за чување података

3.3.1. До ког периода ће подаци бити чувани у репозиторијуму? _____

3.3.2. Да ли ће подаци бити депоновани под шифром? Да Не

3.3.3. Да ли ће шифра бити доступна одређеном кругу истраживача? Да Не

3.3.4. Да ли се подаци морају уклонити из отвореног приступа после извесног времена?

Да Не

Образложити

4. Безбедност података и заштита поверљивих информација

Овај одељак МОРА бити попуњен ако ваши подаци укључују личне податке који се односе на учеснике у истраживању. За друга истраживања треба такође размотрити заштиту и сигурност података.

4.1 Формални стандарди за сигурност информација/података

Истраживачи који спроводе испитивања с људима морају да се придржавају Закона о заштити података о личности (https://www.paragraf.rs/propisi/zakon_o_zastiti_podataka_o_licnosti.html) и одговарајућег институционалног кодекса о академском интегритету.

4.1.2. Да ли је истраживање одобрено од стране етичке комисије? Да **Не**

Ако је одговор Да, навести датум и назив етичке комисије која је одобрила истраживање

4.1.2. Да ли подаци укључују личне податке учесника у истраживању? Да **Не**

Ако је одговор да, наведите на који начин сте осигурали поверљивост и сигурност информација везаних за испитанике:

- а) Подаци нису у отвореном приступу
 - б) Подаци су анонимизирани
 - ц) Остало, навести шта
-
-

5. Доступност података

5.1. Подаци ће бити

а) јавно доступни

б) доступни само уском кругу истраживача у одређеној научној области

ц) затворени

Ако су подаци доступни само уском кругу истраживача, навести под којим условима могу да их користе:

Приступ подацима може бити омогућен истраживачима за потребе научноистраживачког рада, уз претходни договор са аутором података и институцијом у којој су настали подаци, у складу са условима под којима су подаци добијени и коришћени.

Ако су подаци доступни само уском кругу истраживача, навести на који начин могу приступити подацима:

Захтев за приступ подацима подноси се аутору података, уз навођење сврхе коришћења. Приступ се одобрава након провере услова за коришћење података и у договору са институцијом у којој су подаци настали.

5.4. Навести лиценцу под којом ће прикупљени подаци бити архивирани.

6. Улоге и одговорност

6.1. Навести име и презиме и мејл адресу власника (аутора) података

Оливера Лазих, olja.hrnjakovic@uns.ac.rs - ауторка прилагођеног истраживачког скупа података

коришћеног у оквиру докторског истраживања. Изворни подаци добијени су од Института за ратарство и повртарство, Нови Сад, након чега је извршен одабир, издвајање и прилагођавање података релевантних за предмет и циљеве истраживања.

6.2. Навести име и презиме и мејл адресу особе која одржава матрицу с подацима

Оливера Лазић, olja.hrnjakovic@uns.ac.rs

6.3. Навести име и презиме и мејл адресу особе која омогућује приступ подацима другим истраживачима

Оливера Лазић, olja.hrnjakovic@uns.ac.rs