

# Sustav za snimanje radarom sa sintetičkom aperturom i integriranom klasifikacijom objekata

---

**Turčinović, Filip**

**Doctoral thesis / Disertacija**

**2024**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

*Permanent link / Trajna poveznica:* <https://urn.nsk.hr/urn:nbn:hr:168:296160>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2025-01-10**



*Repository / Repozitorij:*

[FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repository](#)





Sveučilište u Zagrebu  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Filip Turčinović

**SUSTAV ZA SNIMANJE RADAROM SA  
SINTETIČKOM APERTUROM I INTEGRIRANOM  
KLASIFIKACIJOM OBJEKATA**

DOKTORSKI RAD

Zagreb, 2024.



Sveučilište u Zagrebu  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Filip Turčinović

**SUSTAV ZA SNIMANJE RADAROM SA  
SINTETIČKOM APERTUROM I INTEGRIRANOM  
KLASIFIKACIJOM OBJEKATA**

DOKTORSKI RAD

Mentor: Izv. prof. dr. sc. Marko Bosiljevac

Zagreb, 2024.



University of Zagreb

FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Filip Turčinović

**SYSTEM FOR SYNTHETIC APERTURE RADAR  
IMAGING WITH INTEGRATED OBJECT  
CLASSIFICATION**

DOCTORAL THESIS

Supervisor: Associate professor Marko Bosiljevac, PhD

Zagreb, 2024

Doktorski rad izrađen je na Sveučilištu u Zagrebu Fakultetu elektrotehnike i računarstva,  
na Zavodu za komunikacijske i svemirske tehnologije.

Mentor: Izv. prof. dr. sc. Marko Bosiljevac

Doktorski rad ima: 89 stranica

Doktorski rad br.: \_\_\_\_\_

## O mentoru

Marko Bosiljevac rođen je u Zagrebu, gdje je diplomirao i doktorirao na Fakultetu elektrotehnike i računarstva (FER) Sveučilišta u Zagrebu 2005. tj. 2012. godine. Od ožujka 2006. godine radi na Zavodu za radiokomunikacije FER-a (od 2019. Zavod za komunikacijske i svemirske tehnologije); 2006.-2012. kao znanstveni novak, 2012.-2016. kao viši asistent, 2016.-2021. kao docent i od 2021. kao izvanredni profesor. Za iznimno uspješnu doktorsku disertaciju i znanstveni rad 2012. je nagrađen Srebrnom plaketom „Josip Lončar“. Od 2016. godine član je ERASMUS povjerenstva na FER-u i koordinator za studij Informacijskih i komunikacijskih tehnologija za sve programe razmjene studenata. Također je izabrani član Vijeća za tehničke znanosti Sveučilišta u Zagrebu.

Njegovi istraživački interesi uključuju razvoj novih struktura elektromagnetskog prijenosa, antena i metapovršina, razvoj autonomnih sustava za daljinska istraživanja temeljenih na zemaljskim radarima sa sintetičkom aperturom te razvoj različitih svjetlovodnih senzora. U ovom području sudjeluje u nizu znanstveno-istraživačkih projekata te je trenutno i voditelj jednog projekta Hrvatske zaklade za znanost. Također, od 2018. kontinuirano surađuje s KTH Stockholom i UC3M Madrid na razvoju periodičkih struktura s kliznom simetrijom te je od 2019. član upravljačkog odbora COST projekta CA18223 „Future communications with higher-symmetric engineered artificial materials“. Autor je ili koautor više od 100 publikacija u znanstvenim časopisima i zbornicima konferencija, kao i nekoliko poglavlja u knjigama i drugih članaka, patenata i prezentacija. Zbog uspješnog djelovanja u području optike bio je jedan od osnivača Laboratorija za primijenjenu optiku (AOLab) na FER-u.

Aktivno sudjeluje u aktivnostima Hrvatske sekcije IEEE, a posebno Odjela za antene i širenje elektromagnetskih valova (AP) čijem radu je doprinio i kao predsjednik i dopredsjednik odjela u nekoliko mandata. Također, sudjelovao je u organizaciji konferencija ICECom od 2005. do 2023. godine i član je organizacijskih odbora doktorskih škola u sklopu „European School of Antennas“ pod nazivom „Advanced Mathematics for Antenna Analysis“ i „Metasurfaces for Antennas“ te pomoćni urednik u znanstvenom časopisu „Automatika“.

---

## About the Supervisor

Marko Bosiljevac was born in Zagreb, where he graduated and received his PhD at the Faculty of Electrical Engineering and Computing (FER) of the University of Zagreb in 2005 and 2012, respectively. Since March 2006, he has been working at the Department for Wireless Communications of FER (since 2019, Department for Communication and Space Technologies); 2006-2012. as a PhD researcher, 2012-2016. as senior assistant, 2016-2021. as assistant professor and from 2021 as associate professor. In 2012, he was awarded the "Josip Lončar" Silver Plaque for his very successful doctoral thesis and scientific work. Since 2016, he has been a member of the ERASMUS committee at FER and coordinator for the study of Information and Communication Technologies for all student exchange programs. He is also an elected member of the University of Zagreb Council for Technical Sciences.

His research interests include the development of new electromagnetic transmission structures, antennas and metasurfaces, the development of autonomous remote sensing systems based on ground-based synthetic aperture radars, and the development of various optical sensors. In this field, he participates in a number of scientific research projects and is currently the principal investigator of one project of the Croatian Science Foundation. Also, since 2018, he has been continuously collaborating with KTH Stockholm and UC3M Madrid on the development of periodic structures with glide symmetry, and since 2019 he is a member of the steering committee of the COST project CA18223 "Future communications with higher-symmetric engineered artificial materials". He is the author or co-author of more than 100 publications in scientific journals and conference proceedings, as well as several book chapters and other articles, patents and presentations. Because of his successful work in the field of optics, he was one of the founders of the Applied Optics Laboratory (AOLab) at FER.

He actively participates in the activities of the IEEE, especially, Croatian Chapter on Antennas and Propagation to which he contributed as president and vice-president in several mandates. He also participated in the organization of the ICECom conferences from 2005 to 2023 and is a member of the organizational committees of the doctoral schools within the "European School of Antennas" under the names "Advanced Mathematics for Antenna Analysis" and "Metasurfaces for Antennas" and an assistant editor in the scientific journal "Automatika".

## Zahvala

U govoru nakon obrane ovog rada rekao sam da ću zahvale odraditi kroz usmene privatne sesije. Ipak, jedan pisac upozorava da „*ono što nije zapisano, i ne postoji*“ pa ću - zapisati - neka imena koja su me usmjerila prema ovom trenutku.

Hvala kolegama sa ZKIST-a. Više puta sam rekao da je to najbolji zavod na FER-u, a pojedini djelatnici tog Zavoda odlučili su mi dodatno potvrditi tu tezu u tjednima prije obrane.

Hvala ekipi koju inače predstavljam kao cimere. Samo je jedan od njih zadržao tu titulu do danas. Neki to već dugo nisu, neki tek koji mjesec, a neki nisu nikad ni bili. Ono što im je zajedničko je da su svi prerasli razinu cimerstva. Abecednim redom: But, Duraz, Gid, Jajo, Kače, Kosir, Luj, Mikula, Miro, Mori i Načelnik.

Hvala članovima Instituta za promociju znanosti. Posebno DŽ-u koji trenutno piše doktorat, Šumadinu s PMF-a, Ivani i Renati s Kognitivne, Mati s podcasta 'O zakonima prirode' i doc. dr. sc. Kolaru s VERNa. Hvala i ravnatelju Instituta, voditelju Podkista, suputniku i supatniku. Bez njega ne bih bio ovo što sam danas: gori, drukčiji, srozan.

Hvala mojoj zagrebačkoj obitelji - Branimiru, Petri, Marti i Boni - na putovanjima, večerama i blagoslovima.

Hvala Marku čiji bi stil mentoriranja trebao pisati u udžbenicima. Rijetko se dogodi takav spoj znanja, sposobnosti, dobrote i smirenosti.

Završavam zahvalu s onima od kojih u mom životu sve kreće. Hvala mojoj obitelji: mami, tati, Franu i Lauri. Doktorat posvećujem baki, dedi, noni i nonetu.



## Sažetak

Disertacija prati razvoj sustava za snimanje radarom sa sintetičkom aperturom i integriranom klasifikacijom objekata. Sustav je implementiran oko mikroročunala Raspberry Pi što mu omogućuje objedinjavanje radarskog snimanja, rekonstrukcije radarske slike i modela dubokog učenja. Radarsko snimanje radi se po principu zemaljskog radara sa sintetičkom aperturom (eng. *Ground Based Synthetic Aperture Radar*, GBSAR). Obradom snimljenih sirovih radarskih podataka rekonstruira se dvodimenzionalna slika prostora algoritmom Omega-K. Za klasifikaciju objekata temeljenu na tim podacima uspoređeno je nekoliko arhitektura modela od kojih se posebno istaknula arhitektura rezidualne mreže (eng. *Residual Network*, ResNet). Analizirani su pristupi razvoja klasifikacijskih modela treniranih sirovim radarskim podacima i onih treniranih rekonstruiranim slikama. Pokazano je da, iako je ljudima jednostavnije interpretirati snimljenu scenu na rekonstruiranoj slici, pristup koji koristi neobrađene (sirove) podatke postiže bolje rezultate. Za taj je pristup razvijena prilagodba inačice rezidualne mreže s 18 slojeva (ResNet-18) u kojoj se, zbog specifičnog izgleda matrice dobivene nakon radarskog snimanja, ne mijenja veličina njezine horizontalne dimenzije za vrijeme treniranja modela. Istražena je metoda povećanja energetske učinkovitosti sustava komunikacijom između radara i klasifikacijskog modela. Metoda uključuje dva seta snimanja - u prvom su korišteni optimalni radarski parametri, a detaljnije su snimljeni oni objekti za koje model nije uspio sa sigurnošću odrediti klasu. U eksperimentalnom testiranju samo 27% primjera je detaljnije snimljeno, a zadržana je visoka razina točnosti klasifikacije. Analizirana je mogućnost dodavanja novog izvora informacije modelu kroz snimanje scene korištenjem dvije linearne polarizacije. Opisano je i implementirano nekoliko pristupa povezivanja matrica dobivenih radarskim snimanjem horizontalnom i vertikalnom polarizacijom, a predstavljena je i arhitektura modela koji prima matrice odvojeno i za svaku trenira zasebnu granu čiji se vektori značajki spajaju prije potpuno povezanog sloja u jedinstveni izlaz. Kroz više je iteracija pokazano da se korištenjem dodatne informacije u obliku nove matrice snimljene drugom polarizacijom može povećati točnost modela. U sklopu doktorskog rada kreirano je i javno objavljeno pet skupova radarskih podataka.

**Ključne riječi:** radar sa sintetičkom aperturom, klasifikacija objekata, ResNet-18, rekonstrukcija radarske slike, algoritam Omega-K, polarizacija elektromagnetskog vala

---

## Extended Abstract

### System for synthetic aperture radar imaging with integrated object classification

Synthetic Aperture Radar (SAR) systems, unlike optical cameras, provide imaging through clouds and in low-light conditions by transmitting longer wavelengths. The primary challenge with longer wavelengths is low resolution. SAR addresses this by capturing signals from multiple positions along the satellite's trajectory, which synthetically increases the antenna aperture and thereby improves resolution. Ground-Based SAR (GBSAR) follows a similar concept where the sensor that emits and receives millimeter-waves moves along a guide rail on the ground. Such approach has several advantages over satellite-based SAR. GBSARs provide the flexibility for specific applications by directing radiation towards areas of interest and adjusting sensor frequency to better detect monitored processes. Some parts of the frequency spectrum, unusable in satellite SAR due to atmospheric water molecules, do not present a limitation for GBSAR. Therefore, higher frequencies can be used, resulting in better resolution. Additionally, the shorter distance between the radar and the observed area also improves resolution of the GBSAR systems. An important aspect of data acquisition is the repetition period, the time required for the radar to recapture the same area. For most SAR satellites, such period is expressed in weeks, whereas for GBSAR it can be in seconds. Short repetition periods are crucial for situations requiring rapid response, such as monitoring bridges, dams, or landslides. As in various SAR applications, deep learning models are extensively used with GBSAR data, and typical scenarios cover prediction of terrain shifts, structural changes, and classify surfaces and objects.

This thesis presents a GBSAR system for object classification named GBSAR-Pi. The system is developed around a Raspberry Pi microcomputer, enabling multiple functions: radar scanning, image reconstruction, and object classification. During radar scanning, the sensor emits and receives chirp signals with a frequency that linearly changes over time (Frequency Modulated Continuous Wave, FMCW) around a central frequency of 24 GHz. In the implemented system the sensor is Innosent IVS-362 FMCW module with integrated voltage-controlled oscillator, transmitting and receiving antennas and mixer. The FMCW module gives a signal obtained after mixing the emitted and received chirps, and whose frequency is proportional to the time delay between them. Measurements from each sensor position along the guide rail are stored on the microcomputer. After completing the scanning process, GBSAR-Pi reconstructs the radar image of the captured area and classifies the objects in the observed scene. The thesis details the implementation of the Omega-K image reconstruction algorithm, which transfers data to the frequency domain to correct irregularities caused by the module's radiation pattern, the linear increase of the emitted signal, and the discrete movement steps of the sensor. This process reconstructs a two-dimensional radar image of the observed area that is easier for humans to

---

interpret. However, the reconstruction process involves approximations and data interpolation, which can distort the integrity of the raw data. To analyze this distortion, two object classification approaches were compared: one where models are trained on raw, unprocessed radar data, and the other where models are trained on reconstructed radar images. Both approaches were tested on various deep learning model architectures suitable for computers with limited resources, such as a single fully connected layer architecture, an LSTM (Long Short-Term Memory) architecture, and Convolutional Neural Network (CNN) architectures like MobileNetV3 and Residual Network (ResNet). The latter in its 18 layer version showed significant promise, leading to its adaptation for raw radar data.

Given that raw data is stored in matrix form, where each column represents a signal obtained after mixing the emitted and received chirp signals, its dimensions are the number of frequency samples on the vertical axis and the number of positions from which the GBSAR system scanned the scene on the horizontal axis. The number of frequency samples, determined by the FMCW module, is 1024, while the number of positions depends on the set radar parameters and is typically between 10 and 120 in our case. Since in standard ResNet18 the input data is halved in five subsampling steps, just before fully connected layer the horizontal dimension of our input matrix is reduced to very small values (often to 1). The idea behind subsampling is to reduce the input dimensions of usually large tensors with three channels, enhancing the efficiency of the training process and generalization. However, due to the significant difference in the dimensions of the recorded GBSAR matrix, which is a relatively small two-dimensional matrix (tensor) of depth 1, we removed the downsampling of the smaller dimension, which in our case represents the number of GBSAR steps. This modification allows the classification model to utilize the information potentially carried by the relationship between FMCW measurements. In comparison with the aforementioned architectures (one fully connected layer, LSTM, MobileNetV3, and standard ResNet18), the model based on modified ResNet18 achieved the highest mean average precision (mAP).

The classification accuracy of the model trained on reconstructed images (IMG model), based on the standard ResNet18 architecture, was compared to the accuracy of the model trained on raw radar data (RAW model), implemented using the modified architecture. Experimental results show that the RAW model consistently outperformed the IMG model. Specifically, the RAW model misclassified eight examples, whereas the IMG model misclassified 11 out of 67 test set examples. Interestingly, the overlap of misclassified examples is peculiar: no example was correctly classified by the IMG model and misclassified by the RAW model. These findings imply that while reconstructed images are more interpretable to humans, especially regarding the locations of the objects in the scene, the reconstruction process does not introduce new information about the observed object. On the contrary, the process distorts the data and decreases the performance of the model.

---

In the next step, the impact of radar parameters on the classification accuracy of the RAW model was analyzed. By altering radar parameters, the resolution was degraded in both azimuth and range directions. For the former, the number of positions from which the scene was observed was reduced while maintaining the total aperture, while for the latter, the frequency bandwidth of the emitted chirp was decreased. Specifically, we experimented with five GBSAR step cases (the scene was captured from 100, 50, 33, 17, and 14 positions) and three bandwidth cases (1300 MHz, 600 MHz, and 300 MHz). Results showed that classification accuracy decreased in both cases. Such behavior is expected since lowering radar resolution in any direction reduces the quantity of information about the observed object. It is important to note that each data point in the datasets captured from a smaller number of positions was artificially generated by excluding FMCW measurements from the original matrix. Hence, such datasets were bigger than the original one, which enabled the analysis of models trained using more examples. These results indicated that high accuracy can be achieved with GBSAR data captured from fewer positions if the dataset is big enough, prompting the idea of energy-efficient system based on two classification processes.

The first model is trained with matrices acquired using fewer GBSAR steps, and the second using matrices representing the same scene but captured from additional positions. In the experimental setup, the first matrix was obtained using 10 GBSAR steps, while in the second case, the scene was observed from 30 additional positions, resulting in matrices of dimensions  $1024 \times 10$  and  $1024 \times 40$ , respectively. The system initially classified objects based on the first dataset and repeated the measurement process using additional positions for parts of the scene where the model could not classify objects with high enough confidence (in our case, over 80%). Model confidence was determined using the Softmax function, which normalizes model outputs into a probability distribution. Given the changed Softmax function task, its results were evaluated by dividing predicted probabilities into groups and analyzing whether the frequency of correct predictions matched the probabilities. For example, if all cases where the Softmax function determined a probability between 70% and 80% were correctly classified in 70% to 80% of cases, the function for that group was evaluated as correct. Results of the evaluation confirmed the function's accuracy for all groups. Regarding the results of the energy efficiency experiment, even when GBSAR scanned the object with only 10 steps, the classification accuracy achieved over 80% in all five conducted iterations. However, when the system repeated the measurements for the objects that were not classified with satisfactory confidence in the initial classification, the final accuracy reached over 90%. Out of a total of 270 test examples, 73 needed to be re-scanned from additional positions. Hence, the system saved energy by scanning 73% of the examples with four times fewer GBSAR steps than usual while still achieving comparable classification results.

The problem of reducing the number of GBSAR steps correlates with the reduced quan-

---

tity of information carried by the output matrix. That quantity can be increased by adding new information sources. The thesis describes the possibility of using different polarizations of electromagnetic waves emitted by the sensor during scanning to capture the same scene in multiple ways, i.e., with multiple polarizations, as a new source of information to the model. Differences in data captured with horizontal (HH) and vertical (VV) polarizations were analyzed, and the results of models trained on them were compared. Since models were not equally accurate in classifying test objects, it can be concluded that data captured with different polarizations carry different information. Besides the original two datasets — one containing matrices of raw radar data captured with horizontal polarization and the other with vertical polarization — several additional datasets were generated by combining elements of the originals:

- SUB set: matrices generated by subtracting VV matrix from HH matrix.
- AVG set: matrices generated by averaging elements of HH and VV matrices.
- MIX\_ROWS set: matrices generated by alternating rows of HH and VV matrices.
- MIX\_COLS set: matrices generated by alternating columns of HH and VV matrices.
- JOIN set: matrices generated by joining the two matrices along the horizontal axis.

In the first two datasets, the matrix dimensions are the same as those with horizontal (HH) or vertical (VV) polarization, while in the last three, the number of elements is doubled. In MIX\_ROWS, the vertical dimension is doubled, while in MIX\_COLS and JOIN, the horizontal dimension is doubled. Each dataset was used to train, validate, and test a separate classification model based on the previously mentioned adapted residual network architecture. Results showed that the HH model more accurately classified in all iterations compared to the VV model, which can be attributed to the appearance and position of the observed objects and the used FMCW module, whose radiation pattern depends on polarization. All models trained on additional datasets, except for the JOIN approach, performed worse than the HH model. These results can be observed from a perspective wider than classification based on radar data. They highlight the importance of the element order within the matrix on classification results using the same architecture. In the JOIN dataset, the element order is the same as in the original matrices, retaining information about signal positions and arrangements, which is not the case with MIX\_COLS and MIX\_ROWS. In the next iteration, more complex models were tested. The first was an Ensemble model, where Softmax values at the output were calculated as the average of the Softmax outputs of the HH and VV models. The second was a Siamese model, where HH and VV matrices were separately input to train two branches each based on the modified ResNet18 architecture. The branches joined their feature vectors into one common feature vector followed by a fully connected layer that produced a single output. Interestingly, even though these were more complex models, they yielded comparable results to the JOIN model. Such an outcome was attributed to the fact that the models were trained with small datasets. However, the finding of the superiority of the JOIN approach is notable. The problem

---

of dataset size for deep learning models regarding radar data repeatedly emerged in literature and conducted experiments, which is why all datasets created during these analyses are publicly available.

To conclude, an energy-efficient GBSAR-Pi system was implemented, featuring a proposed operation mode. This mode consists of two scene captures: the first with a reduced number of GBSAR steps, and, if needed, a second capture using additional steps. The necessity for the repeated capture is determined by the Softmax function, with a confidence threshold set at 80% in our example setup. We demonstrated that simple communication between the deep learning model and radar can provide energy savings by adapting radar parameters based on model results. During the scanning process, the GBSAR system's sensor emits and receives EM waves with both horizontal and vertical polarizations. The system then combines the two output matrices using the JOIN method, which achieved the highest classification accuracy in all conducted experiments. The classification model, implemented using a modified ResNet18 architecture, is based on raw radar data, as this approach outperformed the image-based method in the comparison of raw and reconstructed data.

Since the thesis consists of two components (radar imaging and object classification), there are two theoretical chapters: remote sensing and deep learning. The first chapter describes the basics of radar and chirp signals, with particular emphasis on SAR. It then details the mathematical model of the Omega-K reconstruction algorithm, illustrating its implementation with images generated at each phase. The chapter concludes with a description of the GBSAR-Pi system implementation. The next theoretical chapter introduces CNNs, covering the basics of their architectures and the principles of the CNN training process. It briefly outlines the architectures of the recurrent neural network LSTM and MobileNetV3, while providing a detailed description of the residual network architecture, which is the basis for all later implemented models. The chapter concludes with an overview of the metrics used to evaluate the models. The next three chapters present the results and contributions of the doctoral dissertation. The first of these chapters describes five radar datasets that differ in the observed scenes and the radar parameters used in the scanning process. The second chapter presents the results of analyses conducted on models trained with these datasets. Notably, it includes a subchapter comparing classification based on raw data versus reconstructed images, as well as a subchapter on adding information to the model using two linear polarizations. The final chapter covers the integration of GBSAR and object classification into a single comprehensive system. The conclusion highlights the most significant contributions of the doctoral dissertation.

**Keywords:** Ground Based Synthetic Aperture Radar, Object Classification, ResNet18, Radar Image Reconstruction, Omega-K Algorithm, EM Wave Polarization

# Sadržaj

<b>1. Uvod</b>	1
1.1. Pregled literature	.4
<b>2. Daljinska istraživanja</b>	8
2.1. Osnove radara	.9
2.1.1. Frekvencijski modulirani signal	.11
2.2. Radar sa sintetičkom aperturom	.13
2.2.1. Rekonstrukcijski algoritam Omega-K	.14
2.2.2. Polarizacija	.19
2.3. Implementacija sustava GBSAR-Pi	.20
2.3.1. Postupak snimanja radarske slike	.22
2.3.2. Implementacija algoritma Omega-K	.23
<b>3. Duboko učenje</b>	28
3.1. Osnove konvolucijske neuronske mreže	.29
3.1.1. Arhitektura mreže	.30
3.1.2. Učenje mreže	.33
3.2. Klasifikacijski modeli	.37
3.2.1. ResNet-18	.37
3.2.2. LSTM i MobileNetV3	.39
3.3. Metrike	.40
<b>4. Kreirani skupovi podataka</b>	42
4.1. Parovi sirovih radarskih podataka i rekonstruiranih slika	.42
4.2. Scene snimljene iza prepreke	.43
4.3. Scene snimljene radarom s različitim parametrima	.44
4.4. Scene s 9 različitih objekata	.46
4.5. Scene snimljene različitim polarizacijama	.46

<b>5. Klasifikacija objekata pomoću sustava GBSAR-Pi</b>	48
5.1. Usporedba pristupa sirovim podacima i rekonstruiranim slikama	.48
5.1.1. Klasifikacijski modeli temeljeni na sirovim radarskim podacima	.49
5.1.2. Modifikacija arhitekture ResNet-18	.51
5.1.3. Klasifikacijski modeli temeljeni na rekonstruiranim slikama	.52
5.1.4. Rezultati usporedbe dvaju pristupa	.52
5.2. Modeli trenirani podacima snimanim različitim radarskim parametrima	.56
5.2.1. Analiza na istom broju primjera	.57
5.2.2. Analiza na drugačijem broju primjera	.58
5.3. Klasifikacija objekata s optimizacijom parametara	.59
5.3.1. Određivanje sigurnosti modela	.62
5.4. Klasifikacija na skupovima podataka s različitim polarizacijama	.63
5.4.1. Rezultati modela trenirani na kreiranim skupovima	.64
5.4.2. Zajednički i Sijamski modeli	.67
5.4.3. Analiza funkcije Softmax	.69
5.4.4. Utjecaj polarizacije na klasifikaciju određenog objekta	.70
<b>6. Sustav za snimanje SAR-om i integriranom klasifikacijom objekata</b>	74
<b>7. Zaključak</b>	76
<b>Literatura</b>	79
<b>Životopis</b>	88
<b>Biography</b>	89



# Poglavlje 1

## Uvod

Zlatno doba daljinskih istraživanja najavljuju dvije velike misije promatranja Zemlje iz svemira koje kreću 2024. godine. NISAR, kao zajednički projekt NASA-e (od eng. *National Aeronautics and Space Administration*) i Indijske svemirske istraživačke organizacije (eng. *Indian Space Research Organization*, ISRO) pratit će procese ekosustava, reljefne promjene i kretanje ledenjaka, dok će se misijom *Biomass* Europske svemirske agencije (eng. *European Space Agency*, ESA) nadzirati šume i njihov utjecaj na klimatski sustav. Sateliti će u oba slučaja koristiti koncept radara sa sintetičkom aperturom (eng. *Synthetic Aperture Radar*, SAR) koji snima radarske slike odašiljanjem i primanjem mikrovalova zadržavajući visoku rezoluciju. Za razliku od optičkih slika koje koriste valne duljine vidljivog dijela spektra (400 – 700 nm), NISAR će Zemlju promatrati u frekvencijskim područjima L i S koji odgovaraju valnim duljinama između 7.5 i 15 cm, a Biomass u frekvencijskom području P čiji su valovi duljine oko 70 cm.

Signali većih valnih duljina mogu dublje prodirati kroz prepreke što se kod slanja signala sa satelita prvenstveno odnosi na mogućnost snimanja kroz oblake, dok se u slučaju Biomassa može spomenuti i prodiranje kroz krošnje drveća. S druge strane, na tako velikim udaljenostima senzori u mikrovalnom području postižu lošiju rezoluciju. Budući da veća apertura antene poboljšava rezoluciju, SAR koristi kretanje satelita sa sensorom po putanji kako bi određeno područje snimio s više pozicija i kombinacijom snimki simulirao veću aperturu. Na sličnom se principu temelji i SAR postavljen na tlo (eng. *Ground Based SAR*, GBSAR) koji pomicanjem senzora po vodilici dugoj nekoliko metara snima prostor ispred sebe kako bi kreirao dvodimenzionalnu radarsku sliku snimljenog područja. Takav pristup u odnosu na satelitski SAR ima nekoliko prednosti. Posebno treba istaknuti fleksibilnost u prilagođavanju radara određenoj primjeni. Osim pozicije GBSAR-a i usmjeravanja zračenja, moguće je prilagoditi frekvenciju senzora kako bi se bolje detektirali procesi koje nadzire. Također, dijelovi frekvencijskog spektra se kod satelitskog SAR-a ne mogu koristiti zbog molekula vode u atmosferi što kod GBSAR-a ne predstavlja ograničenje, a, budući da se radi o višim frekvencijama, to znači da se SAR konceptom na tlu može ostvariti bolja rezolucija. Na rezoluciju utječe i udaljenost

između radara i promatranog prostora koja je, opet, u slučaju GBSAR-a povoljnija. Iz perspektive prikupljanja podataka važno svojstvo je period ponavljanja snimanja, odnosno vrijeme potrebno da radar drugi puta snimi isto područje. To je vrijeme kod većine SAR satelita iskazano u tjednima, dok se kod GBSAR-a može raditi o sekundama. Kratko vrijeme ponavljanja neophodno je u situacijama u kojima je potrebna brza reakcija kao što su nadzor mostova i brana ili praćenje odrona zemlje. U primjenama koje uključuju metode dubokog učenja, modeli trenirani podacima dobivenim GBSAR-om predviđaju reljefne pomake uzvisina i klizišta, strukturne promjene građevina i klasificiraju površine i objekte.

U ovom je radu opisan GBSAR sustav za klasifikaciju objekata nazvan GBSAR-Pi. Sustav je razvijen oko mikroračunala Raspberry Pi što mu omogućuje objedinjavanje radarskog snimanja, rekonstrukcije radarske slike i klasifikaciju objekata. Za vrijeme radarskog snimanja senzor odašilje i prima signale čija se frekvencija linearno mijenja u vremenu (eng. *Frequency Modulated Continuous Wave*, FMCW) oko centralne frekvencije koja u implementiranom sustavu iznosi 24 GHz. Mjerenja senzora sa svake pozicije na vodilici po kojoj se kreće spremaju se na mikroračunalo, a nakon što je dovršen proces snimanja GBSAR-Pi rekonstruira radarsku sliku snimljenog područja. U radu je opisana implementacija algoritma za rekonstrukciju slike Omega-K. Uspoređena su dva pristupa klasifikaciji objekata. U prvom su modeli trenirani sirovim, neprocesiranim radarskim podacima, a u drugom rekonstruiranim radarskim slikama. Oba su pristupa testirana na nekoliko arhitektura modela dubokog učenja od kojih se posebno istaknula arhitektura rezidualne mreže (eng. *Residual Network*, ResNet). Na toj se arhitekturi temelji prilagodba modela za rad sa sirovim radarskim podacima koji su nakon snimanja pohranjeni u obliku matrice. Dimenzije matrice odgovaraju broju frekvencijskih uzoraka (kojima je predstavljen izlazni signal dobiven nakon miješanja poslanog i primljenog FMCW signala) na vertikalnoj osi i broju pozicija s kojih je scena snimljena na horizontalnoj. Prilagodba uključuje uklanjanje koraka poduzorkovanja horizontalne dimenzije za vrijeme treniranja kako bi model mogao koristiti informacije dobivene iz odnosa između izlaznih signala snimljenih s različitih pozicija.

Implementacija GBSAR-a na mikroračunalu omogućila je programsko postavljanje parametara radara. Provedena je analiza utjecaja pojedinog parametra na količinu informacije usporedbom rezultata klasifikacije. Analizirani su parametri koji utječu na radarsku sliku u smjeru udaljenosti i u smjeru azimuta. Za prvi je mijenjana širina frekvencijskog pojasa poslanog signala, a za drugi udaljenost između dvije susjedne pozicije s kojih senzor provodi mjerenje. Očekivano, veća širina pojasa i manja udaljenost između pozicija dali su bolje rezultate. Problem kod manje udaljenosti između pozicija, odnosno koraka GBSAR-Pi sustava, je što takvo snimanje troši više energije. Zato je predložen energetski učinkovitiji način snimanja temeljen na dvije klasifikacije između kojih se prilagođava korak GBSAR-a. Nakon inicijalnog snimanja scene većim korakom, provodi se klasifikacija objekata i određuje sigurnost modela za svako

predviđanje. U slučaju niske sigurnosti, snimanje tog dijela scene se ponavlja s manjim korakom nakon čega slijedi druga klasifikacija čiji je rezultat konačan. Sigurnost modela određena je funkcijom Softmax kojom se inače normalizira izlaz modela u vjerojatnosnu distribuciju.

Kao novi izvor informacija klasifikacijskom modelu testiran je potencijal korištenja različitih polarizacija elektromagnetskih valova koje senzor odašilje za vrijeme snimanja jedne scene. Kreirana su dva originalna skupa sirovih radarskih podataka ('HH' i 'VV') u kojima su scene snimljene prvo horizontalnom (za skup 'HH') pa onda vertikalnom polarizacijom (za skup 'VV'). Uz njih, generirano je pet dodatnih skupova u kojima su primjeri dobiveni različitim kombinacijama parova primjera iz skupova 'HH' i 'VV' i za svaki od njih je implementiran klasifikacijski model kako bi se mogla usporediti točnost klasifikacije. U tim skupovima matrice sadrže iste elemente kao i matrice iz skupova 'HH' i 'VV', ali postavljene na različite pozicije pa se rezultati njihovih usporedbi mogu promatrati iz perspektive šire od klasifikacije na radarskim podacima. Razvijen je i složeniji model ('Sijamski') koji za istu scenu odvojeno prima po jednu matricu iz originalnih skupova i za svaku trenira zasebnu granu temeljenu na spomenutoj prilagođenoj arhitekturi rezidualne mreže. Grane potom spajaju svoje vektore značajki prije potpuno povezanog sloja nakon čega model daje jedinstveni izlaz.

Nastavak rada organiziran je na sljedeći način: nakon uvoda dan je pregled literature s najvažnijim radovima iz područja dubokog učenja i radara sa sintetičkom aperturom. Nakon toga rad je podijeljen u dvije teorijske cjeline: *Daljinska istraživanja* i *Duboko učenje*. U prvoj su opisane osnove radara i FMCW signala, a poseban je naglasak stavljen na radar sa sintetičkom aperturom. Opisan je i izveden Omega-K rekonstrukcijski algoritam kojim su kasnije rekonstruirane radarske slike te je prikazana njegova implementacija praćena slikama iz svake faze. Na kraju poglavlja opisana je implementacija komponente za radarsko snimanje sustava GBSAR-Pi. Sljedeća teorijska cjelina uvodi konvolucijske neuronske mreže (eng. *Convolutional Neural Networks*, CNN), osnove njihovih arhitektura i princip učenja CNN modela. Ukratko su opisane arhitekture povratne neuronske mreže LSTM (*Long Short-Term Memory*) i konvolucijske MobileNetV3, a detaljnije arhitektura rezidualne mreže na kojoj se temelje svi kasnije implementirani modeli. Na kraju poglavlja dane su metrike kojima su modeli evaluirani. Sljedeća tri poglavlja prate rezultate i doprinose doktorske disertacije. U prvom od njih opisano je pet skupova radarskih podataka koji se međusobno razlikuju u promatranim scenama i korištenim parametrima. U drugom su dani rezultati analiza provedenih nad rezultatima modela treniranim tim skupovima podataka. Posljednje poglavlje rezultata pokriva integraciju radara sa sintetičkom aperturom i modela za klasifikaciju objekata u jedan cjeloviti sustav.

## 1.1 Pregled literature

Dva rada objavljena u razmaku od godinu dana predstavljaju prekretnice, u tom trenutku, nepovezanih znanstvenih područja. Ideju 'računala koje misli' opisao je Alan Turing 1950. [1], a koncept radara koji pomicanjem senzora povećava aperturu Carl Wiley 1951. [2]. Razvoj tehnologije kroz sljedećih je nekoliko desetljeća paralelno poboljšavao računala sve efikasnijih procesora, većih memorija i bržih grafičkih procesnih jedinica s jedne strane te sve preciznije radarske senzore i rekonstrukcijske algoritme s druge.

Taj je napredak sa znanstvene strane praćen novim arhitekturama modela umjetne inteligencije od kojih su iz perspektive ovog rada posebno zanimljivi ResNet [3], LSTM [4] i MobileNet [5]. ResNet i MobileNet kao konvolucijske neuronske mreže temelje se na konvolucijskom sloju prvi put iskorištenom u [6], dok je LSTM, kao predstavnik povratne neuronske mreže [7], nasljednik Elmanove mreže [8]. Što se tiče računalne implementacije modela, uz radne okvire [9, 10], osmišljeno je nekoliko optimizacijskih algoritama s ciljem ubrzavanja procesa treniranja mreže, a u radu su istaknuti Adam [11] i AdaGrad [12].

U području daljinskih istraživanja računalni algoritmi češće se povezuju s rekonstrukcijom radarske slike. Podaci snimljeni radarom sa sintetičkom aperturom [13] rekonstruiraju se u frekvencijskoj ili vremenskoj domeni. Algoritmi frekvencijske domene su Range Doppler [14], Chirp Scaling [15] i Omega-K [16], a vremenske Global Backprojection [17] i novija inačica Fast Factorized Backprojection [18]. U njihovim usporedbama [19, 20, 21, 22] autori zaključuju manju složenost algoritama frekvencijske domene i veću preciznost onih vremenske, ali i da odabir rekonstrukcijskog algoritma treba prilagoditi skupljenim radarskim podacima. Tako u [19] ističu algoritam Range Migration na kojem se temelji Omega-K pogodnim za korištenje u slučaju kontrolirane kretnje SAR senzora što je slučaj u radaru sa sintetičkom aperturom koji je postavljen na tlu (Ground Based SAR, GBSAR). U [23] autori su usporedili algoritam Omega-K s ostalim algoritmima frekvencijske domene i zaključili kako uz korak Stoltove interpolacije [24] taj algoritam bolje fokusira radarsku sliku u slučaju široke aperture i kod snimanja avionskim SAR-om u kojem je promatrano područje ispred nosa aviona (eng. *high-squint SAR*). Rad [25] predstavlja modifikaciju algoritma Omega-K za primjenu u sustavu GBSAR-a čiji senzor odašilje frekvencijski modulirane kontinuirane signale (FMCW) [26]. U ovom radu je ta inačica algoritma korištena za rekonstrukciju radarske slike i detaljnije je opisan u poglavlju 2.2.1, a njegova implementacija u 2.3.2.

Sustav GBSAR-a sastoji se od senzora i vodilice po kojoj se on kreće. U pravilu se razvoj takvih sustava razlikuje u korištenom senzoru, a najčešće se oni temelje na vektorskom analizatoru mreže (eng. *Vector Network Analyzer*, VNA), softverski definiranom radiju (eng. *Software Defined Radio*, SDR) ili jednom od radarskih modula (npr. FMCW modul). Najveću osjetljivost ima VNA, dok je rješenje s modulom često najpristupačnije. Jedan od sustava izve-

den s VNA-om je LISA [27]. LISA radi u frekvencijskom području C (4-8 GHz) i Ku (12-18 GHz), a mjerenja tog sustava služila su za otkrivanje lavina [28] i praćenje snježnog pokrivača [29]. VNA u svojstvu senzora koristili su i u sustavu UK NERC GBSAR [30] za nadgledanje vegetacije Sveučilišta Tohoku u Japanu [31] i GBSAR-u radio-opservatorija Jicamarca [32] u Peruu kojim su u frekvencijskom području Ku pratili reljefne promjene. Reljefne promjene nadgledane su i GBSAR-om u [33] u kojem je senzor izveden pomoću SDR-a. U tom radu ističu programabilnost SDR-a kao glavnu prednost u odnosu na obične module. Dizajn i implementacija višefrekvencijskog sustava GBSAR-a u kojem SDR odašilje i prima FMCW signale detaljno je opisana u [34]. Kod sustava MIMO-GBSAR [35] umjesto senzora po vodilici se kreće cijeli MIMO (od eng. *Multiple-Input Multiple-Output*) radar [36]. Jedan takav sustav za nadziranje strukture mosta predstavljen je u [37]. Taj sustav radi u frekvencijskom području W (77 GHz) i ima širinu pojasa od 2 GHz. Autori u [38] testirali su ideju korištenja MIMO sustava umjesto GBSAR-a u slučaju kratke aperture. U radu su usporedili sustav LISA s predloženim sustavom MELISSA koji ima 12 odašiljačkih i 12 prijamnih antena međusobno udaljenih za polovicu valne duljine. Zaključili su da je sustav MELISSA jednostavnije implementirati, može brže snimiti scenu i uočiti objekte koji se kreću. U tablici 1.1 dan je popis navedenih GBSAR-ova.

**Tablica 1.1:** Popis razvijenih GBSAR sustava i njihovih osnovnih svojstava.

Naziv	Senzor	Frekvencijski pojas	Rad
LISA	VNA	C, Ku	[27]
UK NERC GBSAR	VNA	L, S, C, X	[30]
Tohoku GBSAR	VNA	L, S, C	[31]
Jicamarca GBSAR	VNA	Ku	[32]
SDR-GB-SAR	SDR	S, C	[33]
SDR-FMCW GBSAR	SDR	-	[34]
MIMO GB-SAR	MIMO	W	[37]

Dva znanstvena područja spomenuta na početku danas su povezani primjenom metoda dubokog učenja na radarskim podacima snimljenim konceptom SAR-a. Jedan od problema koji se pritom istaknuo bio je osigurati dovoljno veliki skup podataka zbog čega se većina prvih klasifikacijskih modela oslanjala na javno dostupan skup MSTAR (od eng. *Moving and stationary target acquisition and recognition*) koji sadrži radarske slike vojnih vozila snimljenih SAR radarom u frekvencijskom području X (8-12 GHz) iz zraka [39, 40]. U radu [41] objavljenom 2005. testirano je nekoliko klasifikacijskih tehnika od kojih je najbolji rezultat ostvario algoritam K najbližih susjeda (eng. *K-Nearest Neighbours*, KNN) [42]. Napretkom algoritama

dubokog učenja, a posebno konvolucijskih neuronskih mreža, modeli tih arhitektura pokazali su se učinkovitijima u raspoznavanju objekata [43]. Skok u točnosti modela dubokog učenja na MSTAR bazi podataka detaljnije je opisan u [44]. U tom su radu usporedili rezultate nekoliko pristupa [45, 46, 47, 48], a kao najbolji su istaknuli [48] koji je modelom temeljenim na dvosmjernoj LSTM mreži dostigao 99.9 % točnosti. Podacima sa satelita Sentinel-1 Europske svemirske agencije (ESA), klasificirana su poljoprivredna zemljišta na jugu Francuske u radu [49]. Budući da taj satelit ima period ponavljanja od 6 dana, korišten je model arhitekture LSTM koji može pratiti vremensku povezanost među podacima. LSTM je korišten i u predviđanju reljefnih promjena uslijed odrona [50]. Radarske slike istog satelita u kombinaciji s optičkim multispektralnim slikama snimljenim satelitom Sentinel-2 koriste se u uklanjanju oblaka s optičkih multispektralnih slika. Autori u [51] i [52] za taj su zadatak koristili generativne suparničke mreže (eng. *Generative Adversarial Networks*, GAN) [53] koji se sastoje od dva modela - generatora i diskriminatora od kojih prvi pokušava kreirati primjer ulaznog podatka, a druga razlikovati originalni ulazni primjer od umjetno kreiranoga. Tom metodom su u [54], koristeći isključivo SAR slike, generirali optičke. Podaci snimljeni GBSAR-om kojima su trenirani modeli dubokog učenja zbog visoke rezolucije mogu predvidjeti i nadzirati promjene uzvisina [55, 56] i rudnika [57] te klasificirati vrste terena [58]. Tablica 1.2 prikazuje navedene primjene modela dubokog učenja koji su trenirani podacima dobivenim SAR-om.

Primjetan je manji broj radova u kojem se GBSAR podaci koriste u takvim modelima. Jedan od razloga je nedostatak kvalitetnog i dovoljno velikog skupa podataka snimljenih GBSAR-om za treniranje neuronske mreže. Taj je problem donekle riješen u području satelitskog SAR-a budući da se trenutno provodi nekoliko satelitskih misija koje uključuju SAR snimanja, a čiji su podaci javno dostupni [59, 60]. Iako se radi o sličnim konceptima, takvi se podaci ne mogu izravno koristiti za učenje modela koji se primjenjuje na GBSAR podacima. Zato su sve baze podataka generirane GBSAR-Pi sustavom (koji je detaljno opisan u ovom radu) javno dostupne [61, 62, 63, 64, 65].

**Tablica 1.2:** Primjene modela dubokog učenja nad SAR podacima.

<b>Primjena</b>	<b>Model</b>	<b>Godina objavljivanja</b>	<b>Rad</b>
Klasifikacija	KNN, SVM [66]	2005.	[41]
Klasifikacija	Wishart [67]	2011.	[58]
Klasifikacija	A-ConvNets	2016.	[43]
Klasifikacija	LeNet	2016.	[45]
Klasifikacija	A-ConvNets + AlexNet	2017.	[46]
Klasifikacija	LSTM	2017.	[48]
Modifikacija optičkih slika	GAN	2018.	[51]
Nadzor reljefnih promjena	BPNN [68], SVM ANFIS [69], RNN, RVM [70]	2019.	[57]
Modifikacija optičkih slika	GAN	2020.	[52]
Nadzor reljefnih promjena	T-GCN [71]	2021.	[56]
Generiranje optičkih slika	GAN	2022.	[54]
Nadzor reljefnih promjena	LSTM	2022.	[55]
Nadzor reljefnih promjena	GCN [72] + LSTM	2023.	[50]

## Poglavlje 2

# Daljinska istraživanja

Daljinska istraživanja naziv je za skup metoda kojima se prikupljaju informacije o nekom objektu bez fizičkog kontakta s tim objektom. U užem smislu taj se naziv odnosi na promatranje Zemljine površine koje je započelo u 19. stoljeću fotografijom Pariza iz balona na vrući zrak. Fotografiranje je vjerojatno najpoznatiji primjer metode daljinskih istraživanja u kojoj se prikupljaju valovi dobiveni refleksijom svjetlosti od promatranih objekata. Upravo je korištenje svjetlosti njezin najveći nedostatak jer ju ograničava u slučajevima slabijeg osvjetljenja što je posebno značajno kod promatranja Zemlje sa satelita. U takvim primjenama, uz nemogućnost snimanja noću, dodatni problem predstavljaju prepreke kao što su oblaci, dim i vulkanska prašina kroz koje svjetlost, zbog njezine male valne duljine (između 400 i 700 nm), otežano prodire.

Za razliku od optičkih metoda koje koriste postojeće valove iz prirode, aktivni senzori odašilju generirane valove veće valne duljine kako bi dobili detaljnije informacije o objektima koji ih reflektiraju. Najčešće se radi o elektromagnetskim valovima čija je valna duljina između 1 mm i 1 m što predstavlja frekvencije od 300 GHz do 300 MHz. Uređaje koji odašilju i primaju takve valove po akronimu engleskog pojma *Radio Detection and Ranging* nazivamo radarima. Radar se sastoji od odašiljačke i prijamne antene od kojih prva emitira u prostor energiju u obliku valova koji se potom reflektiraju od objekata, a druga prima dio tih reflektiranih valova. Prvim radarom smatra se onaj za detekciju i lokalizaciju zrakoplova koji je 1935. godine patentirao škotski znanstvenik Robert Watson-Watt. Tom su izumu prethodila istraživanja koja su pokazala postojanje elektromagnetskih valova i opisala njihovo ponašanje, dok se njihova primjena u daljinskim istraživanjima počela konkretnije proučavati za vrijeme Drugog svjetskog rata kada su radari korišteni za noćno praćenje neprijateljskih pozicija.

U poslijeratnim godinama fokus znanstvene zajednice vratio se na promatranje Zemlje iz zraka što je istaknulo problem rezolucije na velikim udaljenostima. Satelit udaljen 800 km koji snima frekvencijom od 10 GHz za rezoluciju od 50 m treba 700 m dugu antenu [73]. Budući da tako nešto nije praktično, Carl Wiley 1951. godine predstavio je metodu koju je prvotno



nazvao *Doppler beam sharpening* [2], a koja je danas poznatija kao sintetička apertura. Radar sa sintetičkom aperturom (eng. *Synthetic Aperture Radar*, SAR) koristi kretanje senzora kako bi umjetno (sintetički) povećao aperturu antene čime se pospješuje rezolucija. Budući da velika udaljenost utječe i na omjer signala i šuma (eng. *Signal to Noise Ratio*, SNR), senzor često odašilje signale u kojima se frekvencija linearno mijenja u vremenu (eng. *Frequency Modulated Continuous Wave*, FMCW) kako bi ih lakše detektirao na prijamu.

Za preciznije praćenje reljefnih promjena, detekciju lavina ili nadzor brana i mostova često se koristi SAR koncept postavljen na tlu (eng. *Ground Based SAR*, GBSAR) koji zbog manje udaljenosti od promatranog područja i veće fleksibilnosti prilikom usmjeravanja može postići bolju rezoluciju. U GBSAR-u senzor se kreće po vodilici dužine nekoliko metara i snima prostor ispred sebe odašiljući signale okomito na smjer kretanja. Signali GBSAR-a mogu biti viših frekvencija nego u slučaju satelitskog SAR-a budući da je radar postavljen na tlo i bliže promatranom području pa je utjecaj oblaka, odnosno molekula vode na prigušenje i udaljenosti na snagu signala smanjen. Mogućnost korištenja viših frekvencija također doprinosi rezoluciji radara, ali i smanjuje njegove fizičke dimenzije.

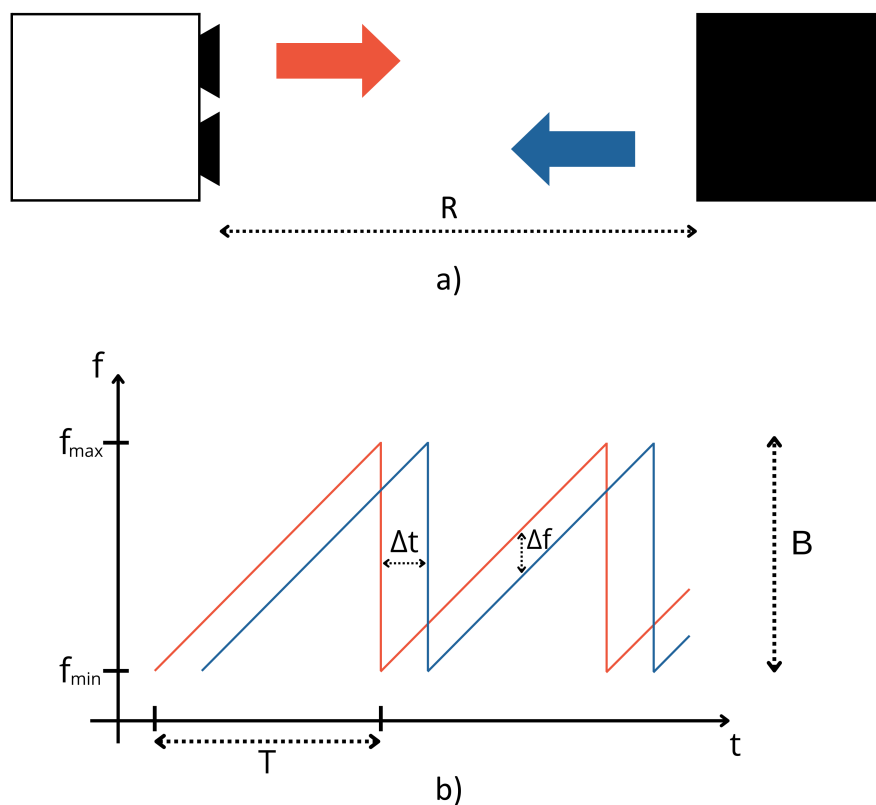
Poglavlje započinje osnovama radara i temeljnim radarskim jednadžbama. U drugom su dijelu opisani frekvencijski modulirani signal koji je odašiljao senzor u razvijenom GBSAR sustavu i sam koncept radara sa sintetičkom aperturom. Naglasak je u nastavku stavljen na algoritam za rekonstrukciju slika koji je primijenjen nad snimljenim radarskim podacima. Nakon toga objašnjeno je svojstvo polarizacije elektromagnetskih valova. Poglavlje završava opisom komponenti sustava GBSAR-Pi za radarsko snimanje i za rekonstrukciju slike. Implementacija rekonstrukcijskog algoritma praćena je grafičkim prikazima po fazama za jedan primjer iz snimljenog skupa podataka.

## 2.1 Osnove radara

Radar je uređaj koji odašilje elektromagnetske valove kako bi temeljem njihovog odjeka detektirao i locirao udaljene objekte. Na primjeru prikazanom na slici 2.1 a), radar odašilje signale prema objektu koji se nalazi na udaljenosti  $R$ . Budući da je poznata brzina širenja vala ( $c = 2.998 \cdot 10^8$  m/s), udaljenost objekta od radara može se odrediti pomoću vremena  $t_d$  potrebnog za put signala od radara do objekta i opet natrag do radara:

$$R = \frac{ct_d}{2}. \quad (2.1)$$

Kako bi se precizno odredilo vrijeme  $t_d$ , snaga primljenog signala mora biti veća od snage šuma. Na tu snagu utječu snaga poslanog signala, dobitak odašiljačke antene, efektivna površina prijamne antene, udaljenost objekta, ali i njegova svojstva. Ta se svojstva u osnovnoj radarskoj



**Slika 2.1:** a) Snimanje objekta radarom. Radar odašilje crveni signal chirp, a prima odbijeni plavi. Objekt se nalazi na udaljenosti  $R$  od radara. b) Izgled poslanog i primljenog chirpa. Na  $y$  osi je frekvencija, a na  $x$  osi vrijeme.

jednadžbi prikazuju u obliku radarskog presjeka (eng. *Radar Cross Section*) koji označavamo  $\sigma$ . Primljenu snagu tada možemo opisati s:

$$P_r = \frac{P_t G}{4\pi R^2} \frac{\sigma A_{ef}}{4\pi R^2} \quad (2.2)$$

gdje su  $P_t$  snaga poslanog signala,  $G$  dobitak odašiljačke antene i  $A_{ef}$  efektivna površina prijamne antene. Prvi dio jednadžbe predstavlja snagu do trenutka raspršenja valova od objekta, a drugi dio od tog trenutka do prijama. Efektivna površina antene može se prikazati i kao:

$$A_{ef} = \frac{G\lambda^2}{4\pi} \quad (2.3)$$

pa se jednadžba 2.2 može zapisati:

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4}. \quad (2.4)$$

Budući da se često snaga primljenog signala gleda kroz odnos sa šumom, jednadžba se često

piše u obliku:

$$\left(\frac{P_r}{N}\right) = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4 N} \quad (2.5)$$

gdje je  $N$  snaga šuma.

Omjer snage poslanog i snage primljenog signala uz poznate parametre radara, može se iskoristiti za računanje rezolucije radara, odnosno najmanji radarski presjek koji objekt može imati kako bi ga radar s tim parametrima mogao detektirati:

$$\sigma = 4\pi \frac{P_r R^4 \lambda^2}{P_t A_e f}. \quad (2.6)$$

Iz te se jednadžbe vidi da se za bolju rezoluciju radara koja u našem slučaju znači smanjenje radarskog presjeka, može smanjiti valna duljina, odnosno povećati frekvencija, smanjiti udaljenost, povećati izračena snaga ili povećati efektivna površina antene.

U slučaju impulsnog radara koji odašilje impulse trajanja  $\tau$ , ako se dva objekta približavaju jedan drugome, njihove udaljenosti se u jednom trenutku neće moći raspoznati jer će se odziv jednog objekta preklapati s odzivom drugog. Rezolucija u smjeru udaljenosti je prema tome uvjetovana i trajanjem impulsa:

$$\rho_r = \frac{c\tau}{2}. \quad (2.7)$$

S praktične strane ta je činjenica ograničavajuća budući da nije jednostavno generirati impulse kratkog trajanja, a dovoljno velike snage. Velika snaga je potrebna jer, prema izrazu 2.4, ona opada s faktorom  $1/R^4$ . Zato se u mnogim radarskim sustavima koristi frekvencijski modulirani signal - *chirp*.

### 2.1.1 Frekvencijski modulirani signal

Chirp (od eng. *Compressed, High-Resolution Pulse*) je signal konstantne amplitude čija se frekvencija linearno mijenja u vremenu. Promjene frekvencija poslanog (crveni) i primljenog (plavi) chirp signala prikazane su na slici 2.1 b). U slučaju rastuće frekvencije (signal koji zovemo *upchirp*) signal počinje od određene minimalne frekvencije ( $f_{min}$ ) i raste za jednu širinu frekvencijskog pojasa chirpa (eng. *Bandwidth, B*) do maksimalne. Omjer širine pojasa ( $B$ ) i vremena trajanja impulsa ( $T$ ) zovemo nagib chirpa i računamo:

$$\gamma = \frac{B}{T}. \quad (2.8)$$

U slučaju rastuće frekvencije uzlaznog chirpa  $\gamma > 0$ , dok je kod silaznog (frekvencija linearno pada kroz vremenski period  $T$ )  $\gamma < 0$ . Ako fazu signala označimo s  $\Phi$ , kod uzlaznog chirpa

promjena frekvencije je linearna s početkom u  $f_{min}$ :

$$\Phi'(t) = f_{min} + 2\pi\gamma t \quad (2.9)$$

što znači da je maksimalna frekvencija  $f_{max} = f_{min} + 2\pi\gamma T$ , a centralna:

$$f_c = \frac{f_{min} + f_{max}}{2} = f_{min} + \pi\gamma T. \quad (2.10)$$

Faza chirp signala jednaka je integralu promjene kružne frekvencije:

$$\Phi(t) = \int_0^t 2\pi(f_{min} + \gamma t) dt = 2\pi f_{min}t + 2\pi\gamma \frac{t^2}{2} = 2\pi f_{min}t + \pi\gamma t^2 \quad (2.11)$$

pa signal koji FMCW senzor odašilje možemo zapisati u obliku analitičkog signala:

$$S_t(t) = e^{i(2\pi f_{min}t + \pi\gamma t^2)}. \quad (2.12)$$

Signal koji radar prima jednak je poslanom signalu pomaknutom u vremenu pa ga možemo zapisati kao

$$S_r(t) = S_t(t - t_d) \quad (2.13)$$

$$S_r(t) = e^{i(2\pi f_{min}(t-t_d) + \pi\gamma(t-t_d)^2)} \quad (2.14)$$

gdje je  $t_d$  vrijeme potrebno signalu od odašiljača do objekta i nazad. Ako znamo to vrijeme, možemo izračunati i udaljenost objekta od senzora jer je  $t_d = \frac{2R}{c}$ , ali njegovo mjerenje zahtijeva preciznu elektroniku i, u slučaju male udaljenosti, mogućnost mjerenja vremena u nano- ili pikosekundama. Zato se udaljenost računa pomoću razlike u frekvenciji između poslanog i primljenog signala koja je proporcionalna vremenu kašnjenja:

$$\Delta f = \gamma t_d \quad (2.15)$$

pa se udaljenost može iskazati kao

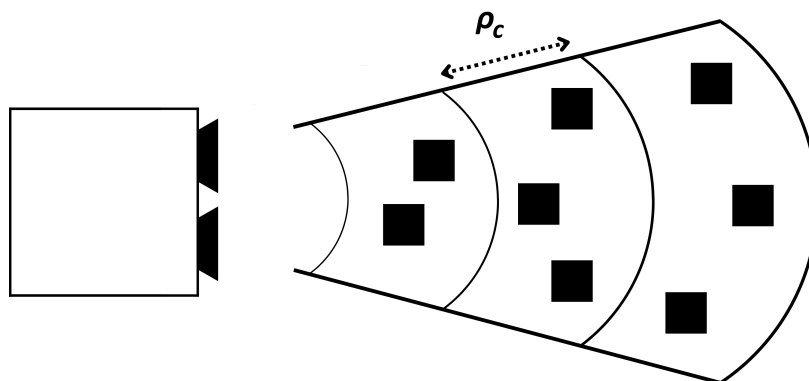
$$R = \frac{cT\Delta f}{2B}. \quad (2.16)$$

Tako generirani impuls lakše je detektirati kod prijama zbog specifičnog ponašanja u vremenu, odnosno zbog linearne promjene frekvencije koja se očekuje i kod reflektiranog impulsa. Kako bi ga se dodatno izdvojilo od šuma, primljeni se signal kroz određenu vremensku instancu uspoređuje s poslanim chirpom. To se radi kompresijom impulsa koja se temelji na metodi prilagođenog filtra čiji je izlaz korelacija primljenog i poslanog signala. Na taj način prilagođeni filter koncentrirajući energiju primljenog signala oko trenutka  $t_d = \frac{2R}{c}$  (očekivano kašnjenje signala reflektiranog od objekta na udaljenosti R) povećava odnos signala i šuma.

Za razliku od rezolucije impulsa konstantne frekvencije koja ovisi o njegovom vremenskom trajanju, rezolucija chirpa ovisi o širini frekvencijskog pojasa:

$$\rho_c = \frac{c}{2B}. \quad (2.17)$$

To znači da razlika između dva objekta u smjeru zračenja mora biti veća od  $\rho_c$  kako bi ih radar koji odašilje chirp signale raspoznao kao dva objekta. Međutim, ako su dva objekta dovoljno udaljena u smjeru koji je okomit na smjer zračenja, ali su nedovoljno udaljeni u smjeru zračenja kao na slici 2.2, njih će radar i dalje vidjeti kao jedan objekt. Statičnim FMCW radarom dobivamo informaciju samo o udaljenosti, ali ne i o lokaciji objekata. Ta se informacija može saznati pomicanjem FMCW radara u prostoru.



**Slika 2.2:** Rezolucija FMCW radara. Objekte koji se nalaze u istoj grupi po udaljenosti (eng. *Range bin*) senzor će detektirati kao jedan objekt. U primjeru sa slike detektirala bi se tri objekta umjesto 8.

## 2.2 Radar sa sintetičkom aperturom

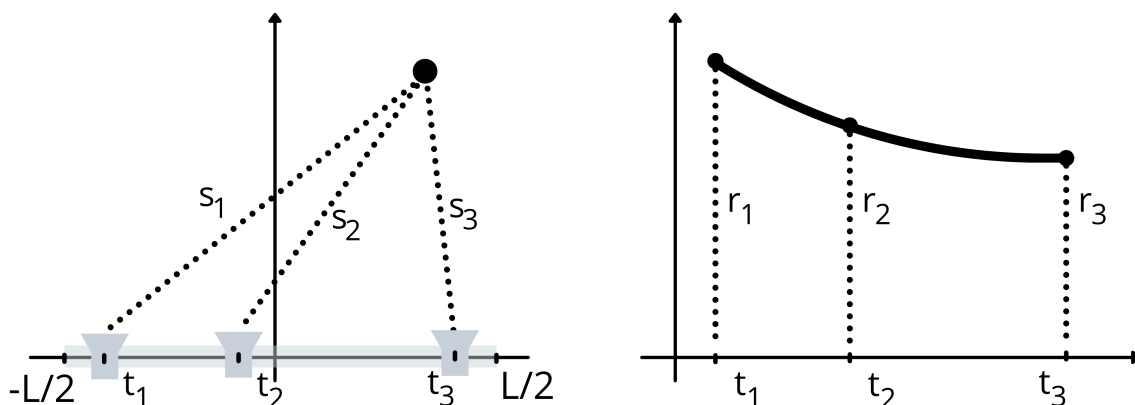
Radar sa sintetičkom aperturom koristi rezultate snimanja FMCW radara, odnosno senzora u SAR sustavu, s nekoliko lokacija kako bi kreirao dvodimenzionalnu radarsku sliku promatranog prostora. Pomicanjem senzora SAR imitira skupljanje podataka dugim antenskim nizom kojim se smanjuje širina izračenog snopa i poboljšava rezolucija. Kod satelitskog SAR-a senzor se zajedno sa satelitom kreće po putanji, dok se kod GBSAR-a kreće po vodilici snimajući prostor okomito na smjer kretanja. U oba slučaja senzor s nekoliko pozicija na putu kojim se kreće daje informaciju o udaljenosti do promatranih objekata u prostoru.

Nakon što se prolaskom senzora do kraja putanje dovrši jedno SAR mjerenje, nad skupljenim podacima sa svih pozicija provodi se rekonstrukcijski algoritam kojim se generira radarska slika snimanog prostora koju je jednostavnije interpretirati budući da su udaljenosti prilagođene stvarnom prostoru. Ovisno o domeni u kojoj se obrađuju ti podaci razlikujemo frekvencijske

i vremenske algoritme. Algoritmi vremenske domene su manje osjetljivi na krivudavost trajektorije snimanja i brzinu kretanja senzora, ali su složeniji. Njihova je složenost reda  $O(n^3)$ , dok je složenost algoritama frekvencijske domene  $O(n^2 \log_2(n))$ . Budući da je u našem slučaju kretanje senzora kontrolirano, odnosno senzor se kreće po ravnoj putanji, a za svako se FMCW mjerenje zaustavlja kako bi uklonio utjecaj kretanja za vrijeme snimanja, slike su rekonstruirane algoritmom Omega-K [25] koji podatke obrađuje u frekvencijskoj domeni. Uz njega se u toj domeni najčešće koristi Range Doppler Algoritam (RDA) [14]. RDA u frekvencijskoj domeni po udaljenosti uklanja nepravilnosti dobivene zbog linearnog porasta frekvencije chirpa, zatim u frekvencijskoj domeni po azimutu poravnava udaljenosti za svako mjerenje senzora primjenom kompenzacije Range Cell Migration (RCMC) i na kraju inverznim Fourierovim transformacijama vraća prikaz u prostornu domenu. S druge strane kod algoritma Omega-K prvo se u potpunosti prelazi u (dvodimenzionalnu) frekvencijsku domenu, zatim se, kao i kod RDA, uklanja rezidualna faza do koje je došlo zbog linearnog povećanja chirpa pa se radi kompresija po udaljenosti i azimutu prema referentnoj udaljenosti. Na kraju se Stoltovom interpolacijom [24] rekonstruiraju podaci kako bi se poravnale točke prostora. Nakon toga se, isto kao u slučaju RDA, inverznim Fourierovim transformacijama podaci vraćaju u prostornu domenu. U nastavku je detaljnije opisan algoritam Omega-K.

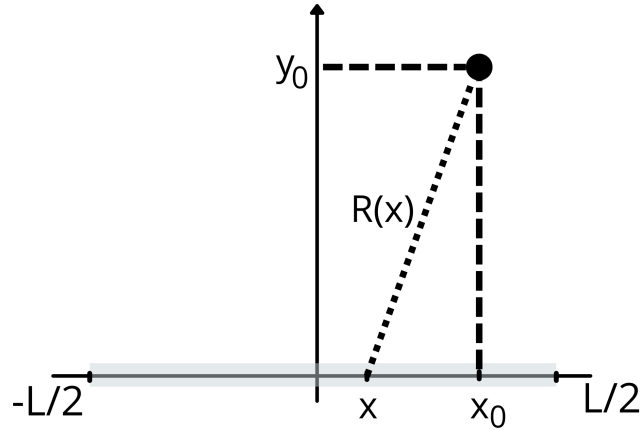
### 2.2.1 Rekonstrukcijski algoritam Omega-K

U pojednostavljenom slučaju prikazanom na slici 2.3 prikazano je snimanje GBSAR-om jedne točke u prostoru nizom FMCW mjerenja. Senzor se kreće po vodilici duljine  $L$  i u trenucima  $t$  snima udaljenosti  $r$ . Na desnom dijelu slike za svaki od trenutaka prikazana je izmjerena udaljenost do promatrane točke.



**Slika 2.3:** Prikaz snimanja jedne točke u prostoru. Na lijevoj slici prikazan je proces snimanja u tri trenutka ( $t_1$ ,  $t_2$  i  $t_3$ ), a na desnoj slici su označene udaljenosti u ta tri trenutka. Ukupan put koji senzor prelazi za vrijeme snimanja je  $L$ .

Cilj algoritma za rekonstrukciju radarske slike je iz tog niza mjerenja generirati sliku koja prikazuje tu točku u prostoru kao što je prikazano na slici 2.4. Položaj senzora u nekom trenutku označen je s  $x$  i može biti između  $-L/2$  i  $L/2$ . Funkcija udaljenosti između senzora na poziciji  $x$  i promatrane točke označena je s  $R(x)$ . Pretpostavka je da se senzor kreće konstantnom brzinom  $v$ .



**Slika 2.4:** Cilj rekonstrukcijskog algoritma je odrediti koordinate  $x_0, y_0$  za promatranu točku.

Na svakoj se poziciji odašilje i prima chirp. Što znači da prije navedeni izraz poslanog chirpa (jednadžba 2.12) zapisujemo kao funkciju dvije varijable:

$$S_t(t, x) = e^{i(2\pi f_{min}t + \pi\gamma t^2)}. \quad (2.18)$$

Isto ponavljamo i s primljenim signalom (jednadžba 2.14) gdje  $t_d$  možemo zamijeniti s  $\frac{2R(x)}{c}$ :

$$S_r(t, x) = e^{i\left(2\pi f_{min}\left(t - \frac{2R(x)}{c}\right) + \pi\gamma\left(t - \frac{2R(x)}{c}\right)^2\right)}. \quad (2.19)$$

Udaljenost na poziciji  $x$  računa se pomoću razlike ta dva signala do koje se dolazi njihovim miješanjem, odnosno množenjem:

$$S_{IF}(t, x) = S_r(t) S_t^*(t) \quad (2.20)$$

$$S_{IF}(t, x) = S_t\left(t - \frac{2R(x)}{c}\right) S_t^*(t) \quad (2.21)$$

$$S_{IF}(t, x) = e^{i\left(2\pi f_{min}\left(t - \frac{2R(x)}{c}\right) + \pi\gamma\left(t - \frac{2R(x)}{c}\right)^2\right)} e^{-i(2\pi f_{min}t + \pi\gamma t^2)} \quad (2.22)$$

$$S_{IF}(t, x) = e^{-i2\pi\frac{2R(x)}{c}f_{min}} e^{-i2\pi\frac{2R(x)}{c}\gamma t} e^{i\pi\gamma\left(\frac{2R(x)}{c}\right)^2}. \quad (2.23)$$

### Rezidualna video faza

Dio oko  $t_d^2$ , odnosno u izrazu 2.23 oko  $\left(\frac{2R(x)}{c}\right)^2$  pojavio se zbog linearnog rasta frekvencije poslanog chirpa u vremenu. Taj dio naziva se rezidualna video faza (eng. *Residual Video Phase*, RVP) i potrebno ga je odstraniti jer negativno utječe na kvalitetu radarske slike nakon rekonstrukcije. Uklanjanje bi se u teoriji moglo učiniti množenjem cijelog prethodnog izraza s negativnom rezidualnom fazom  $e^{-i\pi\gamma\left(\frac{2R(x)}{c}\right)^2}$ . Problem je što u tom trenutku nemamo informaciju o udaljenosti  $R(x)$  između senzora na poziciji  $x$  i promatranog objekta pa ju moramo dobiti iz drugih parametara. Ono što se može iskoristiti je informacija o frekvenciji cijelog signala dobivenog nakon miješanja. Naime, frekvencija cijelog signala može se izračunati pomoću jedinog dijela koji se mijenja u vremenu (dio oko lokalnog vremena  $t$ ), a to je  $e^{-i2\pi\frac{2R(x)}{c}\gamma t}$ . To znači da se cijeli signal mijenja kružnom frekvencijom:

$$\omega = 2\pi \frac{2R(x)}{c} \gamma. \quad (2.24)$$

Frekvenciju signala možemo dobiti kroz izraz  $\omega = 2\pi f$  pa je

$$f = \frac{2R(x)}{c} \gamma. \quad (2.25)$$

Dakle, frekvenciju signala  $S_{IF}$  možemo dobiti primjenom Fourierove transformacije na taj signal pa  $\frac{2R(x)}{c}$  prikazujemo kao  $\frac{f}{\gamma}$ . Rezidualnu fazu sada možemo zapisati:

$$S_{RVP} = e^{-i\pi\frac{f^2}{\gamma}}. \quad (2.26)$$

Množenjem  $S_{IF}$  s tim izrazom uklonit ćemo rezidualnu fazu:

$$S(t) = S_{IF} \cdot S_{RVP} \quad (2.27)$$

$$S(t, x) = e^{-i2\pi\frac{2R(x)}{c} f_{min}} e^{-i2\pi\frac{2R(x)}{c} \gamma t} e^{i\pi\gamma\left(\frac{2R(x)}{c}\right)^2} e^{-i\pi\frac{f^2}{\gamma}} \quad (2.28)$$

nakon čega ostaje

$$S(t, x) = e^{-i2\pi\frac{2R(x)}{c} f_{min}} e^{-i2\pi\frac{2R(x)}{c} \gamma t} \quad (2.29)$$

$$S(t, x) = e^{-i\frac{4\pi}{c}(f_{min} + \gamma t)R(x)}. \quad (2.30)$$

### Množenje referentnom funkcijom

U preostalom izrazu udaljenost na poziciji  $x$  možemo zapisati kao  $R(x) = \sqrt{(x_0 - x)^2 + y_0^2}$ , a  $x$  kao  $x = x_n + vt$  budući da se pozicija senzora mijenja i za vrijeme jednog snimanja pa ga pišemo pomoću početnog položaja senzora za  $n$ -to mjerenje  $x_n$  i brzine kretanja senzora  $v$ . Izraz tada



glasi:

$$S(t, x_n) = e^{-i\frac{4\pi}{c}(f_{min} + \gamma t)\sqrt{(x_0 - x)^2 + y_0^2}}. \quad (2.31)$$

Konačni cilj rekonstrukcijskog algoritma je dobiti koordinate točke  $(x_0, y_0)$  što znači da se te dvije varijable u izrazu 2.31 trebaju odvojiti. Prvi korak je prelazak u funkciju po varijablama u spektru  $S(f_t, x_n)$  označavajući  $\gamma t$  kao frekvenciju  $f_t$ :

$$S(f_t, x_n) = e^{-i\frac{4\pi}{c}(f_{min} + f_t)\sqrt{(x_0 - x_n - v\frac{f_t}{\gamma})^2 + y_0^2}}. \quad (2.32)$$

Za  $f_t$  vrijedi

$$\frac{-B}{2} \leq f_t \leq \frac{B}{2} \quad (2.33)$$

jer je  $-\frac{T_s}{2} \leq t \leq \frac{T_s}{2}$  gdje je  $T_s$  vrijeme trajanja jednog chirpa, a  $\gamma = \frac{B}{T_s}$  pa je  $-\gamma\frac{T_s}{2} \leq f_t \leq \gamma\frac{T_s}{2}$ . Sljedeći korak je prijeći u oblik u kojem se ne mijenja frekvencija, već valni broj  $k$  po kojem je algoritam i dobio ime.

$$k = \frac{\omega}{\lambda} = \frac{2\pi f}{c}. \quad (2.34)$$

Uvodimo oznake valnih brojeva za frekvencije  $f_{min}$  i  $f_t$ :

$$K_{rc} = \frac{4\pi f_{min}}{c} \quad (2.35)$$

$$\Delta K_r = \frac{4\pi f_t}{c} \quad (2.36)$$

i njihov zbroj:

$$K_r = K_{rc} + \Delta K_r. \quad (2.37)$$

$S(f_t, x_n)$  tako postaje  $S(K_r, x_n)$ :

$$S(K_r, x_n) = e^{-iK_r\sqrt{(x_0 - x_n - v\frac{c\Delta K_r}{4\pi\gamma})^2 + y_0^2}}. \quad (2.38)$$

Kako bismo u potpunosti prešli u domenu valnog broja,  $x_n$  moramo predstaviti valnim brojem, a to možemo učiniti Fourierovom transformacijom:

$$S(K_r, K_x) = \int_{-\infty}^{\infty} e^{-iK_x x_n} S(K_r, x_n) dx_n \quad (2.39)$$

$$S(K_r, K_x) = \int_{-\infty}^{\infty} e^{-iK_x x_n} \cdot e^{-iK_r\sqrt{(x_0 - x_n - v\frac{c\Delta K_r}{4\pi\gamma})^2 + y_0^2}} dx_n \quad (2.40)$$

$$S(K_r, K_x) = \int_{-\infty}^{\infty} e^{i\left(-K_x x_n - K_r\sqrt{(x_0 - x_n - v\frac{c\Delta K_r}{4\pi\gamma})^2 + y_0^2}\right)} dx_n = \int_{-\infty}^{\infty} e^{i\phi(x_n)} dx_n. \quad (2.41)$$

Budući da za velike promjene  $x_n$  signal  $S$  jako oscilira, kako bismo riješili integral koristimo metodu stacionarne faze (MSF) [74].

Funkcija  $\phi(x_n)$  predstavlja fazu signala koja glasi:

$$\phi(x_n) = -K_x x_n - K_r \sqrt{\left(x_0 - x_n - v \frac{c\Delta K_r}{4\pi\gamma}\right)^2 + y_0^2}. \quad (2.42)$$

U metodi stacionarne faze računamo stacionarne točke funkcije  $\phi$  koje dobijemo izjednačavanjem derivacije faze s nulom:

$$\frac{\partial \phi}{\partial x_n} = 0 \quad (2.43)$$

$$K_r \left(x_0 - x_n - v \frac{c\Delta K_r}{4\pi\gamma}\right) \frac{1}{\sqrt{\left(x_0 - x_n - v \frac{c\Delta K_r}{4\pi\gamma}\right)^2 + y_0^2}} - K_x = 0. \quad (2.44)$$

Radi preglednosti taj izraz možemo zapisati uz  $a = x_0 - v \frac{c\Delta K_r}{4\pi\gamma}$

$$K_r (a - x_n) \frac{1}{\sqrt{(a - x_n)^2 + y_0^2}} - K_x = 0. \quad (2.45)$$

Rješavanjem dobivamo da su stacionarne točke

$$x_{n,1,2}^* = x_0 - vt \mp \frac{K_x y_0}{\sqrt{K_r^2 - K_x^2}}. \quad (2.46)$$

U nastavku radimo s jednom jer se za drugu može pokazati da daje rješenja koja se nalaze izvan prostora koji smo snimili radarom:

$$x_n^* = x_0 - vt - \frac{K_x y_0}{\sqrt{K_r^2 - K_x^2}}. \quad (2.47)$$

Tu stacionarnu točku prema metodi stacionarne faze uvršavamo u drugu derivaciju faze i u  $\phi(x_n)$  kako bismo dobili konačan izraz:

$$\frac{\partial^2 \phi}{\partial x_n^2} = \frac{-K_r y_0^2}{\sqrt[3]{(a - x_n)^2 + y_0^2}} \quad (2.48)$$

$$\phi(x_n^*) = -\sqrt{K_r^2 - K_x^2} y_0 - K_x x_0 + \frac{vc\Delta K_r K_x}{4\pi\gamma} \quad (2.49)$$

$$S(K_r, K_x) = \sqrt{\frac{2}{|\phi''(x_n^*)|}} e^{i\phi(x_n^*)} e^{-i\frac{\pi}{4}}. \quad (2.50)$$

Radi jednostavnosti pišemo  $A = \sqrt{\frac{2}{|\phi''(x_n^*)|}}$  pa cijeli izraz možemo zapisati kao:

$$S(K_r, K_x) = A e^{-i\sqrt{K_r^2 - K_x^2} * y_0} e^{-iK_x x_0} e^{i\frac{vc\Delta K_r K_x}{4\pi\gamma}} e^{-i\frac{\pi}{4}}. \quad (2.51)$$

Došli smo do izraza u kojem su  $x_0$  i  $y_0$  razdvojeni.

U sljedećem koraku želimo ukloniti treći dio izraza  $\left(\frac{vc\Delta K_r K_x}{4\pi\gamma}\right)$  koji predstavlja fazni pomak zbog pomicanja radara u diskretnim koracima. To činimo množenjem  $S(K_r, K_x)$  s referentnom funkcijom koja sadrži i referentnu udaljenost  $R_s$ , odnosno udaljenost oko koje želimo fokusirati radarsku sliku:

$$S_{RFM} = e^{i(-R_s K_r + R_s \sqrt{K_r^2 - K_x^2})} \quad (2.52)$$

$$S(K_r, K_x) \cdot S_{RFM} = A e^{-i\sqrt{K_r^2 - K_x^2} * y_0} e^{-iK_x x_0}. \quad (2.53)$$

### Stoltova transformacija

Kako bismo preciznije prikazali lokaciju promatrane točke na rekonstruiranoj radarskoj slici, trebamo prijeći iz koordinatnog sustava  $(K_r, K_x)$  u  $(K_x, K_y)$ . To radimo Stoltovom transformacijom koja interpolira podatke iz  $K_r$  u jednako raspoređene podatke u  $K_y$ :

$$K_y = \sqrt{K_r^2 - K_x^2} \quad (2.54)$$

$$S(K_x, K_y) = A e^{-iK_y y_0} e^{-iK_x x_0} e^{i\frac{vc\Delta K_r K_x}{4\pi\gamma}} e^{-i\frac{\pi}{4}}. \quad (2.55)$$

To znači da izraz možemo prikazati u obliku

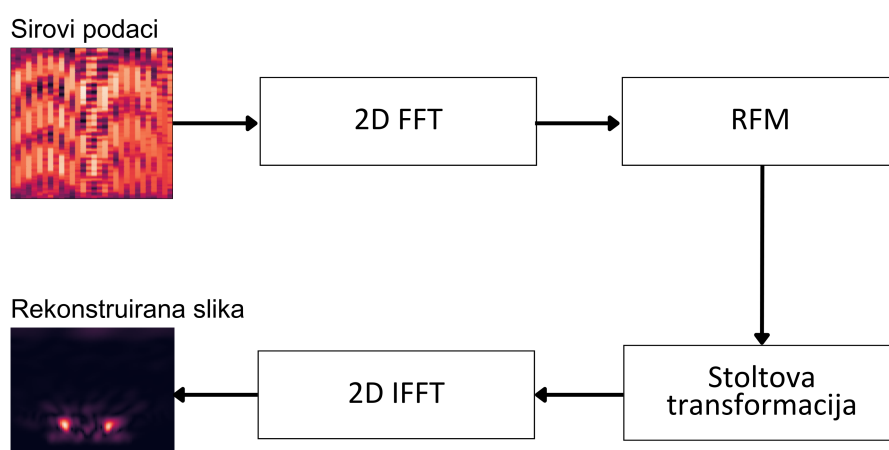
$$S(K_x, K_y) = A e^{-iK_x x_0} e^{-iK_y y_0} \quad (2.56)$$

iz kojeg dvodimenzionalnom inverznom Fourierovom transformacijom (2D-IFT) možemo rekonstruirati sliku. Svi koraci algoritma Omega-K su prikazani na slici 2.5.

### 2.2.2 Polarizacija

Valove čiji su titraji okomiti na smjer kretanja nazivamo transverzalnima i oni, za razliku od longitudinalnih imaju dodatno svojstvo – polarizaciju koja opisuje smjer titranja valova. Transverzalni val možemo prikazati u koordinatnom sustavu superpozicijom dviju komponenti. Budući da mi promatramo planarne elektromagnetske valove, komponente predstavljaju električno i magnetsko polje. U tom slučaju ako se val u koordinatnom sustavu kreće po  $z$ -osi, električno i magnetsko polje će imati komponente u smjerovima  $x$  i  $y$  osi. Po konvenciji polarizacija elektromagnetskog vala predstavlja smjer titranja električnog polja.

Kod linearne polarizacije smjer titranja električnog polja kreće se po ravnoj liniji, dok se



**Slika 2.5:** Blok dijagram algoritma Omega-K.

kod kružne i eliptične smjer titranja polja rotira. Linearna se polarizacija dalje dijeli, ovisno o tome oko koje od dviju osi komponente električnog polja titraju, na horizontalnu i vertikalnu. Po konvenciji se smjer titranja oko osi  $x$  naziva horizontalna polarizacija, a oko osi  $y$  vertikalna. S obzirom na odašiljanje i primanje linearno polariziranih valova kod radara razlikujemo četiri polarizacijske kombinacije:

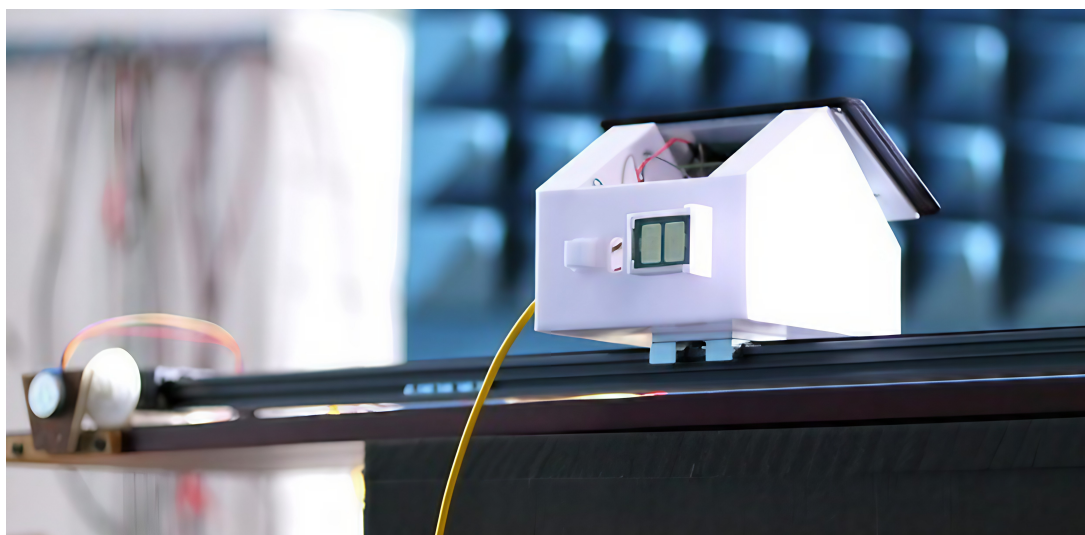
- HH, odašilje i prima horizontalnu polarizaciju
- VV, odašilje i prima vertikalnu polarizaciju
- HV, odašilje horizontalnu, a prima vertikalnu polarizaciju
- VH, odašilje vertikalnu, a prima horizontalnu polarizaciju

Ovisno o sposobnosti radara da odašilje i prima jednu od kombinacija, radar može imati jednostruku (eng. *Single-pol*; HH ili VV), dvostruku (eng. *Dual-pol*; HH, VV, HV ili VH) ili četverostruku polarizaciju (eng. *Quad-pol*) koja može odašiljati i primati horizontalno i vertikalno polarizirane valove istovremeno. U SAR sustavima najčešće se koristi linearna polarizacija, a informacije dobivene različitim polarizacijama primjenjuju se u agrikulturi, šumarstvu, geologiji i oceanografiji [75, 76], dok kod GBSAR-a one omogućavaju nadgledanje okoline i prirodnih fenomena te prevenciju katastrofe [77]. U našem će se slučaju podaci dobiveni različitim polarizacijama koristiti kao dodatna informacija prilikom klasifikacije objekata.

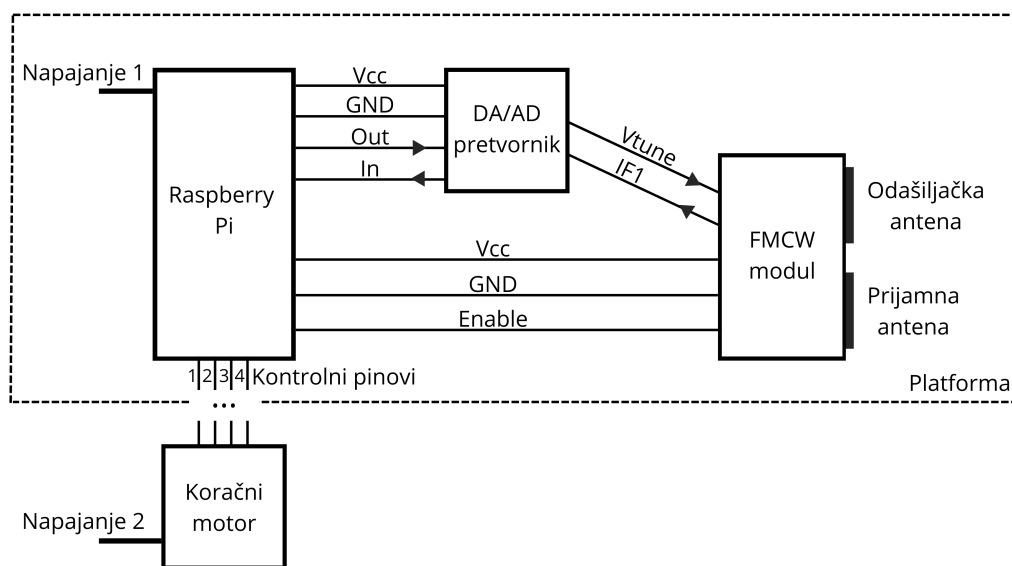
## 2.3 Implementacija sustava GBSAR-Pi

Razvijen je GBSAR sustav za snimanje objekata na malim udaljenostima. Sustav se sastoji od vodilice po kojoj se kreće platforma u kojoj se nalaze senzor, pojačalo i mikroračunalo s analogno-digitalnim (AD) i digitalno-analognim (DA) pretvornikom. Mikroračunalo Raspberry

Pi 4B [78] kontrolira snimanje senzora, obradu i pohranu podataka te kretanje platforme pa je cijeli sustav nazvan GBSAR-Pi. Senzor je FMCW modul Innosent IVS-362 [79] koji ima integrirane odašiljačku i prijamnu antenu, oscilator koji se kontrolira naponom i miješalo, a kretanje platforme po vodilici u stani-kreni načinu rada omogućava koračni motor povezan pomoću četiri kontrolna pina s mikrorračunalom. Bijela platforma sa slike 2.6. integrirana je u mobilni dio vodilice i ima izveden priključak za FMCW modul kojim se može mijenjati polarizacija senzora. Na vrhu platforme je zaslon osjetljiv na dodir kojim se upravlja sustavom.



**Slika 2.6:** Sustav GBSAR-Pi. Mikrorračunalo Raspberry Pi kontrolira FMCW modul koji se vidi na prednjoj strani platforme. Na kraju vodilice na lijevoj strani slike vidi se koračni motor koji pomiče platformu za vrijeme snimanja [80].



**Slika 2.7:** Shematski prikaz implementiranog sustava GBSAR-Pi.

S mikrorračunalom je povezan FMCW modul na način prikazan na slici 2.7. Centralna frekvencija modula je 24 GHz. Osim napona i uzemljenja, prema FMCW modulu idu *Enable* i *V<sub>tune</sub>* pinovi. Prvi se koristi za omogućavanje rada tog FMCW modula i, budući da koristimo samo jedan modul, on je u našoj izvedbi uvijek aktivan. Korištenje tog pina došlo bi više do izražaja u implementaciji GBSAR-a senzorskim nizom. Putem *V<sub>tune</sub>* pina mikrorračunalo preko DA pretvornika generira chirp signal postepeno povećavajući napon koji upravlja oscilatorom na modulu. Prema tehničkim karakteristikama modula napon kojim se određuje minimalna frekvencija ( $f_{min}$ ) je 0.7 V, a maksimalni napon kojim se može postaviti gornja granica frekvencije ( $f_{max}$ ) je 2.5 V s promjenom frekvencije od 720 MHz/V. To znači da je najveća moguća širina frekvencijskog pojasa 1.3 GHz, ali i to da se ta širina može programski podešavati što je iskorišteno za analizu njezina utjecaja na klasifikaciju objekata. Modul ima snagu odašiljanja ekvivalentnu izotropno izračenoj snazi (eng. *Effective Isotropic Radiated Power*, EIRP) 12.7 dBm.

Nakon odašiljanja chirpa, primanja signala istog oblika zakašnjelog u vremenu i miješanja ta dva signala, FMCW modul rezultantski signal prikazan pomoću *In phase (I)* i *Quadrature (Q)* komponente šalje putem *IF1* i *IF2* pinova. Mikrorračunalo u našem slučaju sprema *I* komponentu signala. FMCW modul ima linearnu polarizaciju pa je ona prije svakog mjerenja postavljena zakretanjem kućišta modula na platformi tako da je u mjerenjima korištena ili HH ili VV polarizacija. Kako bi sustav bio u potpunosti mobilan, koriste se dva baterijska izvora napajanja – jedan za koračni motor, a drugi za mikrorračunalo kroz koje se još napajaju AD/DA pretvornik, FMCW modul i zaslon na dodir.

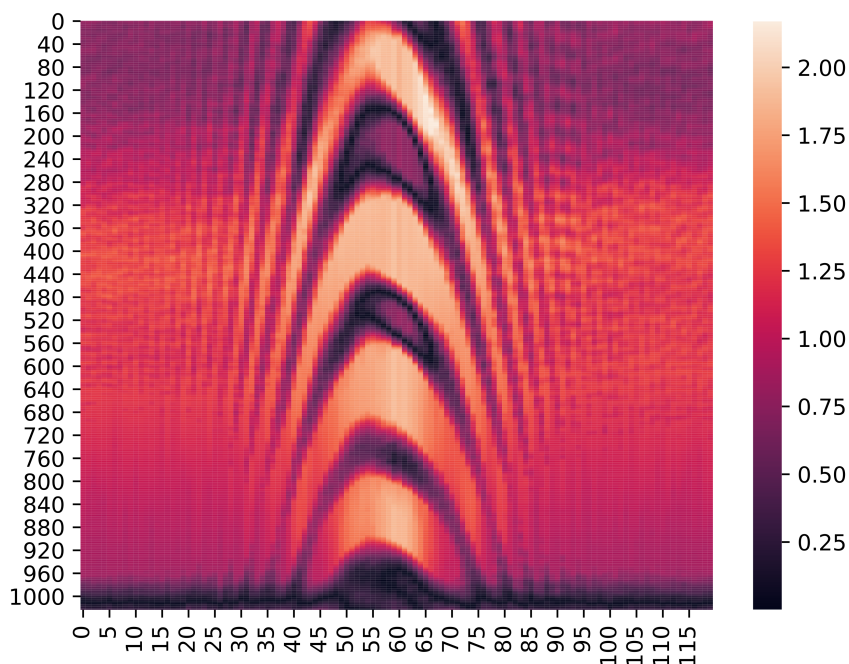
### 2.3.1 Postupak snimanja radarske slike

Mikrorračunalo prilikom svakog mjerenja pokreće skriptu koja prima parametre radara postavljene na grafičkom sučelju. Skripte grafičkog sučelja i samog procesa snimanja pisane su u programskom jeziku *Python*. Radarski parametri koji se mogu postaviti su širina frekvencijskog pojasa, duljina jednog koraka (udaljenost između dva uzastopna mjerenja senzora) i ukupan broj koraka kojim se određuje apertura GBSAR-a. U svakom koraku mikrorračunalo kroz 1024 uzorka kreira rastući napon od 0.7 V do razine određene postavljenom širinom frekvencijskog pojasa (maksimalno 2.5 V). Taj napon generira rastuću frekvenciju chirpa koji se odašilje u promatrani prostor. Nakon miješanja poslanog i primljenog signala chirpa, rezultantski signal određen s 1024 frekvencijska uzorka se sprema u listu koja sadrži trenutne rezultate jednog koraka. U našoj implementaciji lista skuplja rezultate 10 iteracija poslanih i primljenih chirpova s iste pozicije senzora kako bi se njihovim usrednjavanjem povećao odnos signala i šuma. Mikrorračunalo usrednjeni rezultat jednog koraka sprema kao jedan stupac u matrici nakon čega javlja koračnom motoru da pomakne platformu za prije postavljenu duljinu jednog koraka. Koračni motor ima milimetarsku preciznost. Važno je napomenuti da se platforma ne kreće u trenutku

snimanja. Taj se postupak ponavlja dok se scena ne snimi u svim koracima. Rezultantska matrica GBSAR-a čije dimenzije ovise o ukupnom broju koraka (druga dimenzija je 1024) sprema se lokalno na mikroročunalo i na server.

### 2.3.2 Implementacija algoritma Omega-K

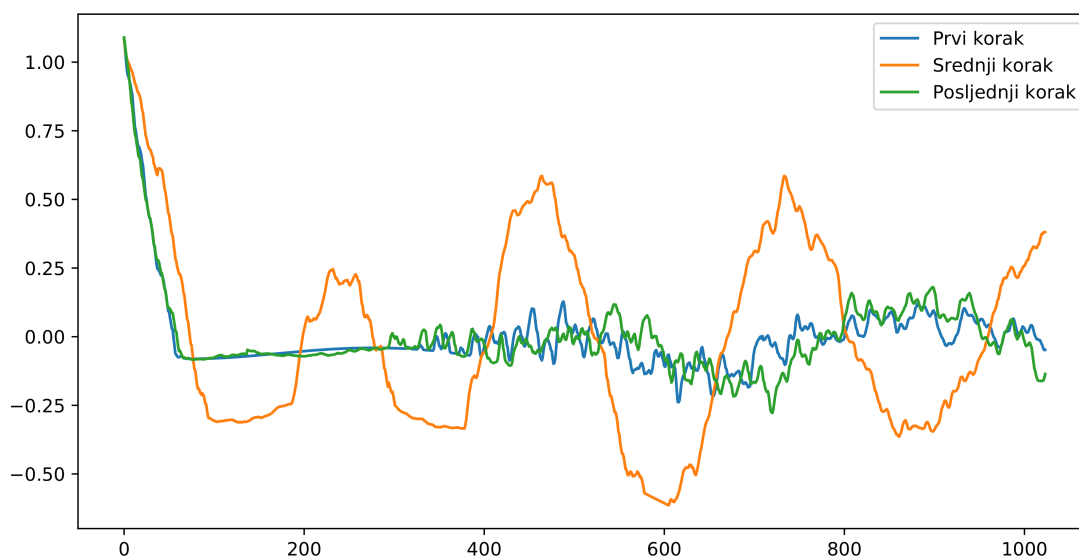
Algoritam Omega-K opisan u 2.2.1 implementiran je u programskom jeziku Python. Skripta koja ga izvodi prima tekstualnu datoteku u kojoj se nalaze rezultati mjerenja iz svakog koraka snimanja GBSAR-om. U snimljenoj sceni na čijem će se primjeru opisati implementacija algoritma postavljen je jedan metalni objekt u sredinu, a scena je snimljena sa 120 koraka duljine 0.5 cm što znači da je ukupna apertura bila 60 cm. Centralna frekvencija poslanog chirpa je 24 GHz, širina frekvencijskog pojasa 1.3 GHz, a vrijeme trajanja perioda jednog chirpa 160 ms. Omjer širine pojasa i trajanja chirpa, odnosno nagib chirpa je u tom slučaju  $\gamma = 8.1 \cdot 10^{10}$ . Postavljena širina pojasa daje rezoluciju u smjeru udaljenosti  $\rho_c = 11.5$  cm. Snimljeni se podaci prvo prebacuju u matricu biblioteke *Numpy* s kojom se mogu dalje obrađivati. Toplinska mapa matrice sirovih podataka u kojima svaki stupac predstavlja mjerenje na jednoj poziciji prikazana je na slici 2.8.



**Slika 2.8:** Toplinska mapa sirovih radarskih podataka snimljenih u 120 pozicija razmaknutih 0.5 cm. U sredini promatrane scene bio je metalni objekt.

Primjeri signala dobivenih nakon mjerenja s prve (1. korak), srednje (60. korak) i posljednje (120. korak) pozicije senzora prikazani su na slici 2.9. Na slici se može primijetiti kako signal

srednje pozicije ima izraženiju nižu frekvenciju u odnosu na ostala dva što je i očekivano budući da je radar na toj poziciji bio najbliže objektu. Prije tog prikaza nad svim je podacima funkcijom *Detrend* biblioteke *Scipy* ublažen utjecaj linearnosti chirpa i linearnom maskom uklonjeno je nekoliko najnižih frekvencija do kojih je došlo zbog preslušavanja između odašiljačke i prijamne antene.



**Slika 2.9:** Prikaz rezultantskih signala snimljenih FMCW senzorom na prvoj, srednjoj i posljednjoj poziciji GBSAR-a. Može se primijetiti veća amplituda i dominantna niža frekvencija signala srednjeg koraka koja implicira manju udaljenost do objekta.

U sljedećem je koraku izvedeno uklanjanje rezidualne faze. Prema jednadžbi 2.28 svaki signal moramo pomnožiti s  $e^{-i\pi\frac{f^2}{\gamma}}$ . U skripti brzom Fourierovom transformacijom (eng. *Fast Fourier Transform*, FFT) biblioteke *Numpy* signal prebacujemo u frekvencijsku domenu i određujemo frekvenciju cijelog sustava  $f$  kako bismo stupce matrice pomnožili s

```
np.exp(1j * np.pi * f**2 / gamma)
```

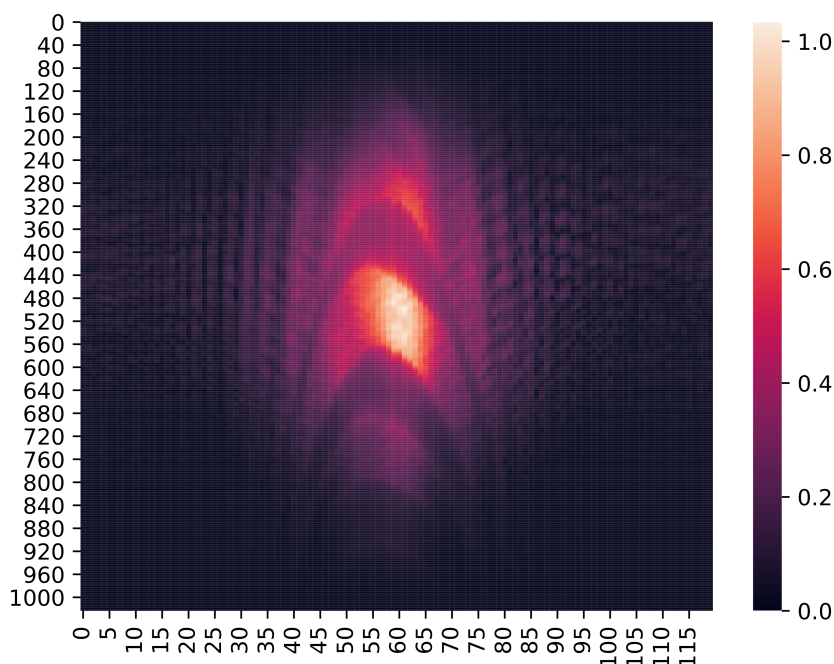
Nakon toga primjenom Hanningovog prozora posebno ističemo one vrijednosti koje nam nose najviše informacije. U našem slučaju to su frekvencije koje predstavljaju udaljenosti na kojima očekujemo objekte. Hanningov prozor množi elemente signala svakog koraka s

$$w(n) = 0.5 - 0.5 \cos\left(\frac{2\pi n}{M-1}\right) \quad (2.57)$$

gdje je  $M$  broj duljina prozora, odnosno kod nas broj uzoraka signala (1024). Slika 2.10 prikazuje toplinsku mapu nakon uklanjanja rezidualne faze i primjene Hanningovog prozora.

Budući da imamo konačan broj koraka (120), prije *FFT*-a po azimutu umjetno dodajemo mjerenja koja ne utječu na konačni rezultat, a poboljšavaju radarsku sliku uglađivanjem rubova





**Slika 2.10:** Toplinska mapa matrice nakon uklanjanja rezidualne faze i primjene Hanningovog prozora.

jer se dodaje cijeli niz uzoraka nakon kojih i rezultantska matrica nakon *FFT*-a ima povećani broj uzoraka. Konkretno, povećali smo broj uzoraka sa 120 na 256, a u svakom od dodanih 136 stupaca nalaze se 1024 nule generirane funkcijom *Zeros* iz biblioteke *Numpy*. Matrica nakon primjene *FFT*-a po azimutu prikazana je na slici 2.11.

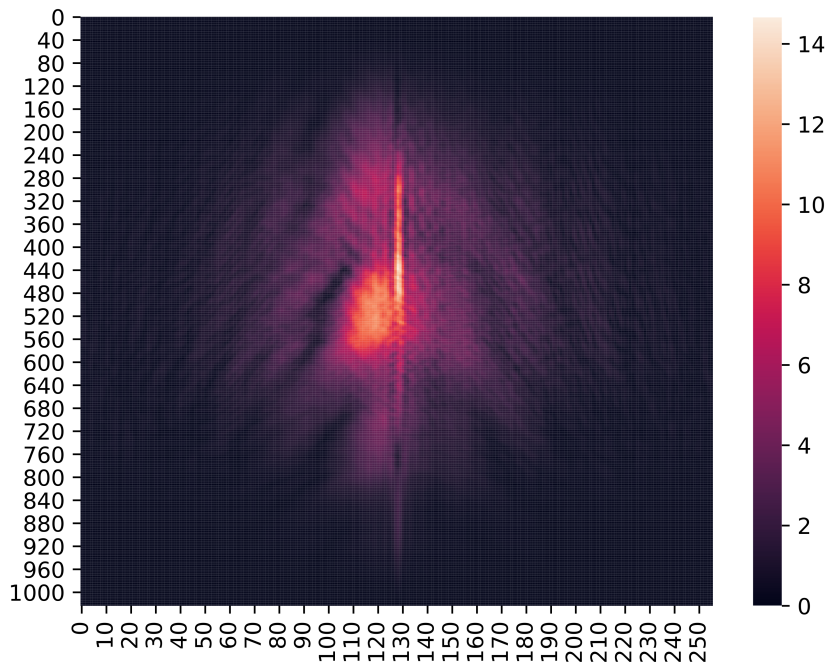
Fourierovom transformacijom po azimutu, prema jednadžbi 2.39 prešli smo u domenu valnog broja u kojoj možemo ukloniti fazni pomak koji se dogodio zbog pomicanja radara u diskretnim koracima pa, referentnu funkciju prikazanu izrazom 2.52 programski uvodimo s:

$$\begin{aligned} \text{phi\_mf}[\text{index\_j}, \text{index\_i}] = & -Rs * \text{kr}[\text{index\_i}] + Rs * (\text{kr}[\text{index\_i}] ** 2 \\ & - \text{kx}[\text{index\_j}] ** 2) ** 0.5 \end{aligned}$$

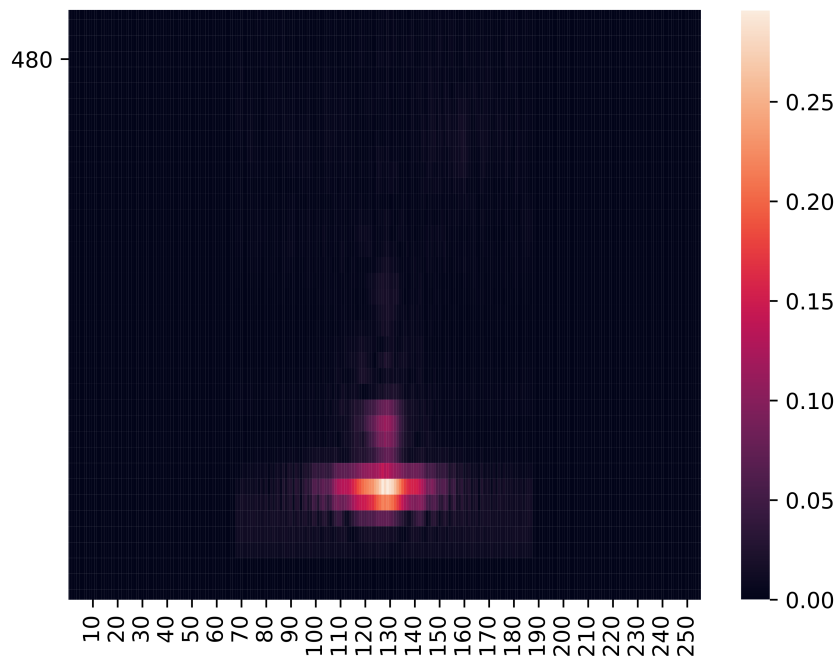
gdje je  $R_s$  referentna udaljenost i u ovom slučaju je postavljena na 0.2 m,  $k_x$  je valni broj po azimutu, a  $k_r$  određujemo prema jednadžbi 2.37, odnosno  $K_r = \frac{4\pi f_{min}}{c} + \frac{4\pi f_t}{c}$  pa programski zapisujemo:

```
kx = np.linspace(-np.pi/step_size, np.pi/step_size, len(data_fft))
kr = np.linspace(((4*np.pi/c)*(fc-b/2)), ((4*np.pi/c)*(fc+b/2)),
                sweep_samples)
```

Nakon toga možemo rekonstruirati radarsku sliku dvostrukom inverznom Fourierovom transformacijom kojom iz domene valnog broja prelazimo u realni prostor. Matrica dobivena funkcijom



**Slika 2.11:** Toplinska mapa matrice nakon dodavanja uzoraka i *FFT*-a po azimutu.



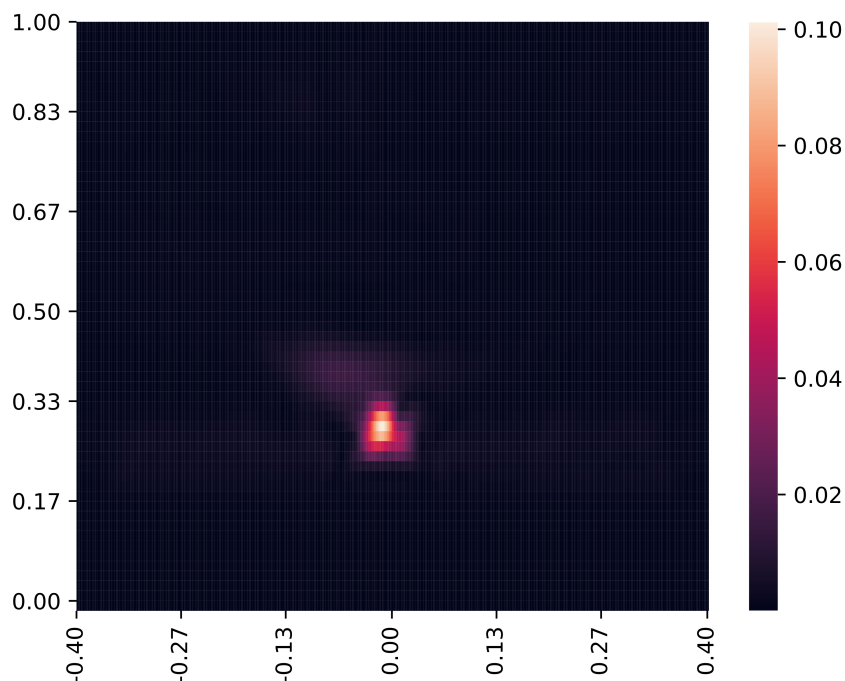
**Slika 2.12:** Toplinska mapa matrice nakon provedenog množenja referentnom funkcijom

*IFFT2* biblioteke *Numpy* nakon množenja referentnom funkcijom (RFM) prikazana je na slici 2.12.

Ipak, budući da smo prije inverzne Fourierove transformacije (IFT) bili u  $K_r$ ,  $K_x$  prostoru, slika 2.12 ne predstavlja prostor u  $x$ ,  $y$  koordinatama. Kako bismo u potpunosti prešli u  $x$ ,  $y$  domenu, prema jednadžbi 2.55 prije IFT-a prvo interpoliramo  $K_r$  prema  $K_y$ . Programski izraz 2.54 pišemo:

```
ky.append((kr**2-kx[count]**2)**0.5)
```

Za samu interpolaciju korištena je funkcija *interpolate.interp1d* biblioteke *Scipy* koja kreira funkciju čije pozivanje aproksimira jednodimenzionalni niz elemenata. Matrica nakon Stoltove transformacije i dvodimenzionalne inverzne Fourierove transformacije prikazana je na slici 2.13. To je ujedno i konačna rekonstruirana radarska slika promatranog prostora. Na generiranoj radarskoj slici vidi se objekt u središtu scene. Usporedbom slika 2.12 i 2.13 vidi se kako Stoltova transformacija u frekvencijskoj domeni prilagođava zakrivljenost realnom prostoru.



**Slika 2.13:** SAR slika promatrane scene u čijem se središtu nalazio metalni objekt. Oznake na  $x$  i  $y$  osi iskazane su u metrima.

## Poglavlje 3

# Duboko učenje

Prema [81] „duboko učenje omogućuje odlučivanje na temelju podataka pronalaženjem i izdvajanjem uzoraka iz velikih skupova podataka koji se precizno preslikavaju u valjane izlazne odluke.“ U računalnom smislu cilj metoda dubokog učenja je pomoću poznatih parova ulaznih i izlaznih podataka stvoriti model koji temeljem samo ulaznih podataka može odrediti izlazne. Zbog preciznosti valja naglasiti kako su tom rečenicom obuhvaćeni samo modeli nadziranog učenja, odnosno modeli u kojima znamo očekivane izlaze za svaki ulazni primjer, a u radu će se predstaviti i opisati isključivo takvi modeli.

Općenitije, duboko je učenje dio strojnog učenja koje je pak dio umjetne inteligencije. Za posljednji se i najopćenitiji termin može reći da proučava izradu *'stroja koji misli'* [82]. Ideju postojanja takvog stroja odbacila je 1842. godine "proročica računalnog doba" Ada Lovelace u svojim bilješkama u kojima je opisala način rada modernog računala za koji je predvidjela kako će moći izvršavati isključivo ono što mu ljudi zadaju, ali neće moći proizvesti nešto samostalno. Na tu se tezu nadovezao Alan Turing u radu „*Computing Machinery and Intelligence*“ (Računalni strojevi i inteligencija) objavljenom više od 100 godina kasnije [1]. Turing je predložio da se umjesto pitanja *'Može li stroj misliti?'*, stroj izazove igrom oponašanja čovjeka, a konačni cilj te igre, koju danas znamo kao Turingov test, jest provjeriti je li umjetna inteligencija dostigla ljudsku razinu takvu da uspije zavarati čovjeka njegovom imitacijom. Ovaj se rad smatra temeljnim u području umjetne inteligencije jer je, između ostalog, predvidio postojanje računala koja će uspješno položiti Turingov test. Danas je jedan od najpoznatijih takvih primjera ChatGPT, sustav koji koristi duboko učenje za obradu prirodnog jezika.

Metode dubokog učenja većinom se zasnivaju na neuronskim mrežama koje koriste ideju da se kombinacijom jednostavnijih modela mogu graditi složeniji na način da se izlaz jednog povezuje s ulazom drugog modela. Neuroni u neuronskoj mreži organizirani su u slojeve koji u pojednostavljenoj podjeli mogu biti ulazni, skriveni i izlazni. Svaki neuron prima određeni broj ulaza s pripadajućim težinama i računa težinski zbroj zbrajanjem produkata ulaza i težina. U drugom koraku neuron propušta težinski zbroj kroz aktivacijsku funkciju koja na njega primje-

njuje nelinearno preslikavanje kako bi mreža mogla bolje modelirati nelinearne odnose. Tako kreirani izlaz neurona prosljeđuje se ili sljedećem neuronu ili predstavlja konačni izlaz modela što znači da se duboko učenje svodi na jako veliki broj množenja, zbrajanja i nelinearnih preslikavanja. Budući da se doista radi o jako velikom broju matematičkih operacija, kako bi se ubrzalo treniranje neuronske mreže koriste se grafičke procesne jedinice prilagođene radu s vektorima i matricama koje takvima promatraju nizove ulaza i izlaza svih neurona.

Posebno zanimljiva primjena neuronskih mreža iz perspektive ovog rada je raspoznavanje slika, odnosno klasifikacija objekata na slikovnim podacima. Za te se primjene najčešće koriste konvolucijske neuronske mreže (eng. *Convolutional Neural Network*, CNN) u kojima su neuroni prilagođeni detekciji lokalnih vizualnih značajki na način da uče funkcije koje prepoznaju njihovu prisutnost, odnosno odsustnost. Na primjeru prepoznavanja rukom pisanih znamenki funkcije nižih slojeva tražile bi kružnice i ravne linije, a dublji slojevi njihove pozicije i odnose. Budući da se te lokalne značajke mogu pojaviti na više mjesta u slici, svojstvo CNN-a koje se posebno istaknulo u ovoj primjeni je translacijska invarijantnost. Sama detekcija vrši se slijednim pretraživanjem određene značajke primjenom iste funkcije na svaki dio slike. Taj se proces naziva konvolucija po kojem je i cijela podvrsta neuronskih mreža dobila ime. Ovisno o broju, vrsti i povezanosti slojeva unutar mreže razlikujemo više arhitektura CNN-a. U ovom radu većina modela temeljit će se na arhitekturi rezidualne mreže (eng. *Residual Network*, ResNet) koja koristi rezidualne blokove koji joj omogućavaju preskakanje slojeva za koje mreža odluči da ne doprinose rezultatu.

U nastavku su opisane osnovne ideje i gradivni blokovi konvolucijskih neuronskih mreža te proces njihovog učenja. Dan je pregled optimizacijskih algoritama s naglaskom na algoritam Adam koji se koristi u ovom radu. Nakon toga opisana su tri modela na čijim se arhitekturama zasnivaju klasifikacijski modeli koji su raspoznavali objekte na radarskim slikama. Na kraju su objašnjene metrike kojima su se evaluirali implementirani modeli.

### 3.1 Osnove konvolucijske neuronske mreže

Konvolucijske neuronske mreže primjenjuju se u računalnom vidu, raspoznavanju na slikama i videu, analizi medicinskih podataka i obradi prirodnog jezika. Budući da se u ovom radu koriste u klasifikaciji objekata na rekonstruiranim radarskim slikama i sirovim radarskim podacima, opisat će se kroz takve primjere. Arhitektura konvolucijske neuronske mreže sastoji se od nekoliko slojeva prilagođenih određenim zadacima. Pojednostavljeni oblik jedne CNN mreže sastoji se od četiri glavna sloja:

- Ulazni sloj predstavlja ulazne podatke. U slučaju slike u RGB formatu s 496x369 piksela, ulazni sloj imat će dimenzije 496x369x3. Ulazni podaci u naše modele koji klasificiraju na rekonstruiranim slikama bit će tih dimenzija, odnosno imat će dubinu 3, dok će u

slučaju sirovih radarskih podataka, budući da se radi o nizu signala, dubina biti 1.

- Konvolucijski sloj nad primljenom matricom primjenjuje konvoluciju određenim brojem filtara, a kao izlaz daje mapu značajki.
- Sloj sažimanja mapira skup prostorno bliskih značajki s ulaza u jednu vrijednost koju prosljeđuje kao izlaz prema sljedećem sloju [82].
- Potpuno povezani sloj sve neurone prethodnog sloja povezuje sa svakim neuronom sljedećeg i kao izlaz daje vektor rezultata modela, odnosno daje konačan rezultat klasifikacije.

U nastavku su detaljnije opisani konvolucijski sloj, sloj sažimanja i potpuno povezani sloj.

### 3.1.1 Arhitektura mreže

#### Konvolucijski sloj

Zadatak konvolucijskog sloja je prepoznavanje značajki u podacima koje prima na ulazu, a glavna metoda kojom to čini i po kojoj je taj sloj, kao i cijela podvrsta neuronskih mreža dobila ime, je konvolucija. U dubokom učenju ona označava proces slijednog prelaska konvolucijskog filtra po ulaznoj matrici u kojem računa izlaz za svaku njegovu lokaciju. Konvolucija se u obradi signala koristi za usporedbu dva signala skalarnim produktom između originala jednog i vremenski pomaknutog i reflektiranog drugog. Na primjeru signala  $x(t)$  i  $h(t)$  konvolucija je dana izrazom:

$$(x * h)(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau. \quad (3.1)$$

U prilagođenom filtru opisanom u poglavlju 2.1.1 koristila se konvolucija poslanog signala i primljenog signala istog oblika (chirpa) kako bi se povećao odnos signala i šuma. U dubokom učenju pod konvolucijom najčešće podrazumijevamo unakrsnu korelaciju koja glasi:

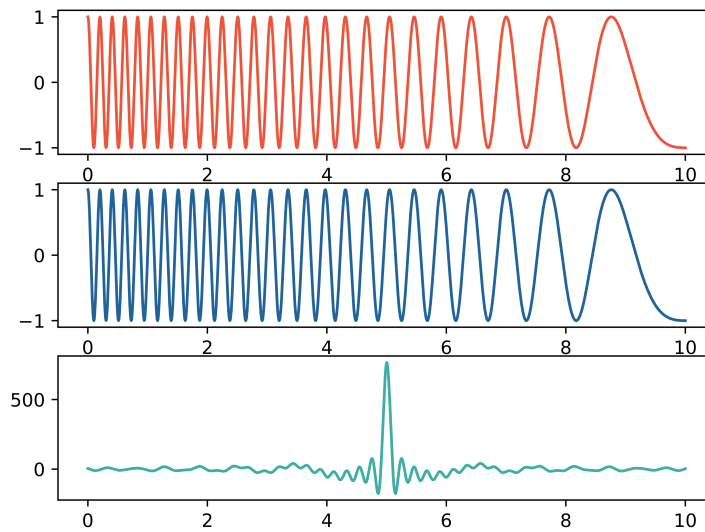
$$(x * h)(t) = \int_{-\infty}^{\infty} x(\tau) h(t + \tau) d\tau. \quad (3.2)$$

Razlog korištenja unakrsne korelacije umjesto konvolucije leži u jednostavnijem shvaćanju procesa. Budući da se vrijednosti filtra postavljaju za vrijeme treniranja mreže, isto je radi li se to na običnoj ili transponiranoj matrici [83]. Na slici 3.1 prikazan je primjer unakrsne korelacije dva jednaka chirp signala.

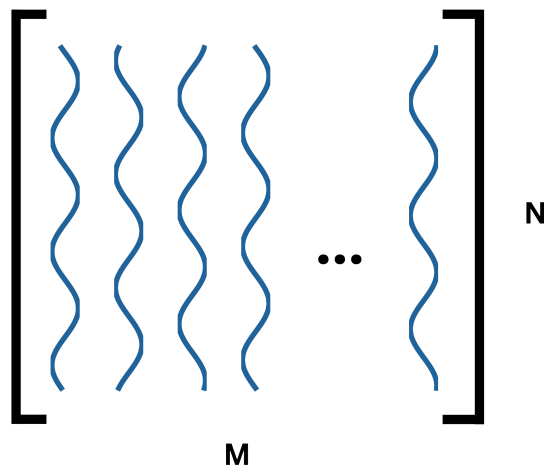
Kod vremenski diskretnih signala unakrsnu korelaciju definiramo kao zbroj svih diskretnih uzoraka:

$$(x * h)(t) = \sum_{\tau=-\infty}^{\infty} x(\tau) h(t + \tau) \quad (3.3)$$

Budući da su u našem slučaju signali dobiveni iz različitih pozicija prikazani u obliku matrice, unakrsnu korelaciju radimo u dvije dimenzije. U sljedećoj jednadžbi niz od  $M$  signala duljine  $N$ , što daje matricu dimenzija  $N \times M$ , predstavlja ulaz  $x(n_H, n_W)$  u kojem je  $n_H$  pojedinačni uzorak



**Slika 3.1:** Dva jednaka chirp signala (crveni i plavi) i njihova unakrsna korelacija.



**Slika 3.2:** Matrica signala dobivenih FMCW radarom s različitih pozicija.

signala, a  $n_W$  određuje sam signal. Odnosno par  $n_H, n_W$  daje vrijednost signala  $n_W$  na mjestu uzorka  $n_H$ . Pojednostavljena skica takve matrice prikazana je na slici 3.2.

Ako 2D filtar dimenzija  $FN \times FM$  označimo s  $f(n_H, n_W)$ , njihovu unakrsnu korelaciju možemo zapisati kao:

$$(x * f)(n_H, n_W) = \sum_{k=0}^{FM-1} \sum_{l=0}^{FN-1} x(n_H + k, n_W + l) f(k, l) \quad (3.4)$$

gdje su  $n_H \leq M - 1, n_W \leq N - 1$  i  $n_H, n_W \geq 0$ .

Na taj način filtar, koji se u literaturi naziva i jezgra (eng. *Kernel*), traži određene uzorke

(značajke) i u izlaznu matricu zapisuje rezultat unakrsne korelacije tog dijela matrice i filtra. Iz perspektive implementacije modela, filter je zapravo matrica s težinama koje se ažuriraju tijekom procesa učenja. U izlaznoj se matrici mogu pronaći pozicije na kojima je tražena značajka uočena pa ju zovemo mapa značajki. Budući da jedan konvolucijski sloj radi s cijelim setom filtera, izlaz neće biti samo jedna mapa značajki, već niz dvodimenzionalnih mapa značajki čiji broj odgovara broju filtera. Taj broj označava i broj kanala izlaza, odnosno dubinu izlaza konvolucijskog sloja.

Na jednostavnom primjeru ulazne matrice dimenzija  $5 \times 5$ , konvolucijski filter dimenzija  $3 \times 3$  može postaviti svoj središnji element u samo 9 elemenata ulazne matrice. To znači da će i izlazna mapa značajki imati 9 elemenata, odnosno biti dimenzija  $3 \times 3$ , ali i to da će rubni elementi ulazne matrice biti pokriveni konvolucijskim filtrom manje puta u odnosu na središnje elemente. Kako bi mapa značajki mogla imati isti broj elemenata kao ulazna matrica, prije procesa konvolucije ulaznim se matricama često dodaju novi stupci i redovi na rubove popunjeni nulama da ne utječu na kasnije izračune. U programskim bibliotekama takve se konvolucije označava sa 'same' (nepromijenjeno) jer mapa značajki zadržava isti broj elemenata kao ulazna matrica, dok se one bez nadopune označava s 'valid' (valjano). Postupak dodavanja nula nazivamo nadopuna (eng. *Padding*). Nadopunu nulama koristili smo i za povećanje broja uzoraka prije Fourierove transformacije u implementaciji algoritma za rekonstrukciju radarskih slika (poglavlje 2.3.2).

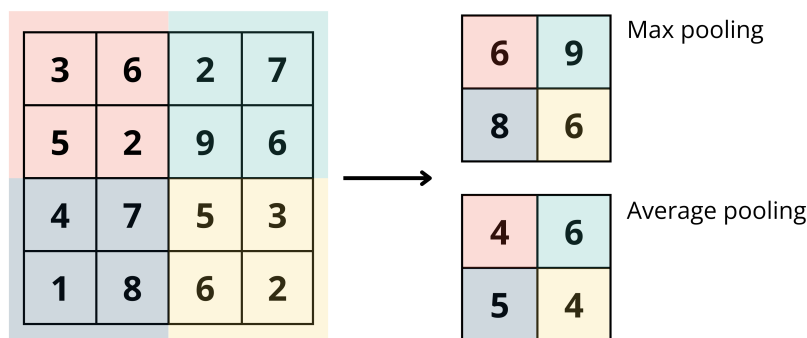
### Sloj sažimanja

CNN koristi sloj sažimanja (eng. *Pooling Layer*) kako bi reducirao broj parametara neuronske mreže i samim tim broj izračuna prilikom učenja, ali i cijeli model učinio robusnijim na manje varijacije pozicija značajki na ulaznoj matrici čime se postiže translacijska invarijantnost spomenuta u uvodu. Taj sloj slijedi nakon konvolucijskog i primjenjuje funkciju sažimanja na mapi značajki kako bi smanjio njezine dimenzije. To se najčešće radi računanjem prosječne vrijednosti elemenata (eng. *Average Pooling*) ili izborom najvećeg elementa (eng. *Max Pooling*) unutar dijela mape značajki određenog veličinom filtra koji se po ulaznoj matrici pomiče korakom  $k$  (eng. *stride*). Na slici 3.3 prikazan je primjer sažimanja izborom najvećeg elementa i prosječnom vrijednosti elemenata korištenjem filtra veličine  $2 \times 2$  i koraka 2.

### Potpuno povezani sloj

Za razliku od konvolucijskog sloja, u potpuno povezanom sloju svi su neuroni povezani sa svim neuronima prethodnog sloja. Taj se sloj najčešće nalazi na kraju konvolucijske mreže i prima jednodimenzionalni niz nakon globalnog sažimanja (eng. *Global Average Pooling*) ili ispravljanja (eng. *Flattening*) višedimenzionalnog tenzora prethodnog sloja, a njegov je izlaz ujedno i izlaz modela i, u slučaju klasifikacijskog modela, predstavlja klase. Potpuna povezanost izlaznih neurona modela omogućava zasebno postavljanje težina, odnosno određivanje utjecaja



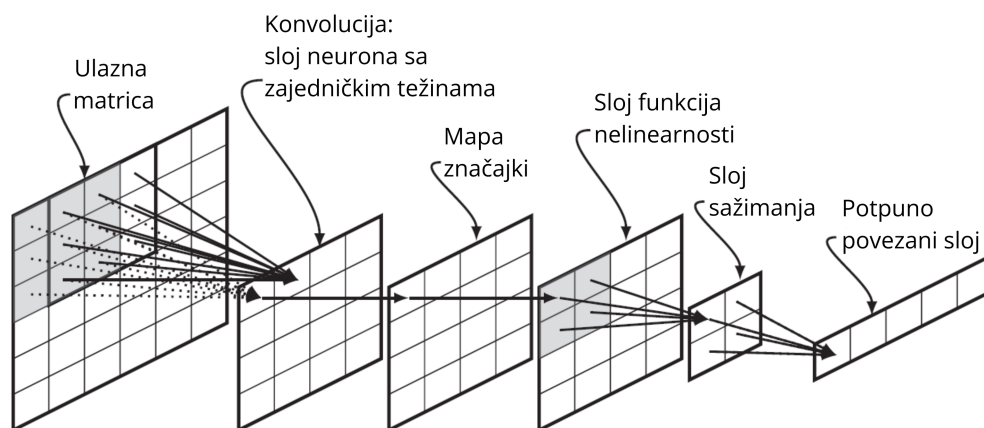


**Slika 3.3:** Primjer sažimanja izborom najvećeg elementa (eng. Max pooling) i prosječnom vrijednosti elemenata (eng. Average pooling) korištenjem filtra veličine 2x2 i koraka 2.

informacija dobivenih različitim jezgrama (filtrima) na pojedinu klasu.

### 3.1.2 Učenje mreže

Primjer jedne jednostavne konvolucijske mreže prikazan je na slici 3.4. Ta se mreža sastoji od ulazne matrice, konvolucijskog sloja koji daje mapu značajki na kojoj se provodi aktivacijska funkcija čiji se izlaz prosljeđuje sloju sažimanja. Svi neuroni tog sloja povezani su sa svim neuronima potpuno povezanog sloja koji daje izlaz modela.



**Slika 3.4:** Primjer konvolucijske mreže koja prima ulaznu matricu, radi konvoluciju i sažimanje te na kraju ima potpuno povezani sloj [81].

Model CNN-a, kao i kod ostalih neuronskih mreža, implementira funkciju koja je određena težinama. Proces učenja neuronske mreže svodi se na traženje skupa težina koji najbolje modelira podatke na kojima uči. Taj se skup težina najčešće traži uspoređivanjem dobivenog i očekivanog rezultata modela i ažuriranjem težina ovisno o tome raste li ili pada razlika između

ta dva rezultata. Proces ažuriranja težina, ili u slučaju konvolucijskog sloja elemenata jezgre, traje dok se ne analiziraju svi podaci iz skupa za treniranje kroz nekoliko iteracija. Algoritam koji se pritom koristi naziva se gradijentni spust. Algoritam gradijentnog spusta prvo računa gubitak za nasumično određene težine i gradijent u toj točki funkcije gubitka. Gradijent je vektor koji je dobiven parcijalnom derivacijom funkcije gubitka po svakoj težini i koji pokazuje smjer najbržeg porasta funkcije u određenoj točki. Nakon toga algoritam „spušta“ promatranu točku po funkciji gubitka u smjeru suprotnom od izračunatog vektora za određeni korak koji se naziva stopa učenja (eng. *Learning rate*). Nova točka u domeni funkcije gubitka označava novi skup težina koje se potom ažuriraju, a algoritam ponovno računa gubitak i gradijent za nove težine te iterativno ponavlja proces dok ne dođe do najmanje vrijednosti gubitka. Na taj način algoritam određuje težine za koje je razlika između rezultata dobivenih modelom i onih očekivanih najmanja, odnosno određuje težine kojima se dani podaci najbolje modeliraju.

Razlikujemo tri načina gradijentnog spusta:

- grupni (eng. *Batch*) u kojem se težine ažuriraju korištenjem svih dostupnih podataka,
- stohastički (eng. *Stochastic*) u kojem se težine ažuriraju temeljem jednog nasumično odabranog primjera i
- mini-grupe (eng. *Mini-Batch*) u kojem se težine ažuriraju temeljem određenog broja nasumično odabranih primjera.

Posljednje navedeni način gradijentnog spusta može se zapisati:

$$\mathbf{w} \leftarrow \mathbf{w} - \eta \sum_{i=1}^N \nabla L(y^{(i)}, h(\mathbf{x}^{(i)}; \mathbf{w})) \quad (3.5)$$

Jednadžba prikazuje promjenu težina  $\mathbf{w}$  gradijentnim spustom.  $\nabla L$  označava gubitak za  $i$ -ti primjer čiji je ulaz  $\mathbf{x}^{(i)}$ , očekivani izlaz  $y^{(i)}$ , a model treniran težinama  $\mathbf{w}$  je za isti ulaz  $\mathbf{x}^{(i)}$  dao izlaz  $h$ . Težine se ažuriraju spuštanjem s razine prošlog skupa težina za vrijednost umnoška stope učenja  $\eta$  i zbroja gubitaka  $N$  primjera iz grupe. Proces ažuriranja težina provodi se unatrag na način da se prvo ažuriraju težine izlaznog sloja, zatim, s informacijom o novim težinama izlaznog sloja, one neposredno prije njih i tako sve do težina prvog skrivenog sloja. Taj se proces stoga naziva algoritam s povratnim postupkom (eng. *Backward propagation algorithm*) [84]. U implementiranim modelima koristio se *Mini-Batch* način gradijentnog spusta.

Optimizacijski algoritam ubrzava proces učenja, odnosno ažuriranja težina spuštanjem gradijenta. Algoritam koji se pokazao učinkovitim na raznim zadacima dubokog učenja je Adam (od eng. *Adaptive Moment Estimation*) [11]. Adam koristi kombinaciju optimizacijskog algoritma RMSProp (od eng. *Root Mean Squared Propagation*) i nešto starije ideje (iz 1964.) korištenja momentuma kojim se ubrzava konvergencija gradijenta u iterativnim metodama [85]. Momentum akumuliranim vrijednostima gradijenta drži generalni smjer pomaka po parametrima čime postaje otporniji na nagle promjene smjera i lokalne optimume [83]. Treba spome-

nuti da se algoritam RMSProp zasniva na još jednom popularnom optimizacijskom algoritmu AdaGrad [12] u kojem se stopa učenja prilagođava za svaki parametar ovisno o tome koliko se taj parametar promijenio tijekom učenja.

S ciljem bržeg učenja često se težine inicijaliziraju predtreniranjem. U tom procesu mreža uči riješiti općenitiji problem inicijalnim postavljanjem težina, a kroz proces treniranja modela se one dodatno podešavaju kako bi se prilagodile konkretnijem problemu. Najčešće se koristi inicijalizacija težina skupom podataka ImageNet [86] koji sadrži u prosjeku 1000 slika za svaku od 100 000 klasa.

## Funkcija gubitka

Funkcija gubitka se u zadacima klasifikacije računa unakrsnom entropijom (eng. *Cross-Entropy Loss*) koja uspoređuje očekivani izlaz s predviđenom vjerojatnošću za taj izlaz iz modela. Za razliku od regresije kod koje se razlike između očekivanih i predviđenih izlaza mogu izravno usporediti zbrojem kvadrata pogreške, kod klasifikacije objekata neki se objekt može ili nalaziti ili ne nalaziti na ulaznom primjeru pa su ti izlazi 1 ili 0. Zato se umjesto binarnog izlaza modela u računanju funkcije gubitka koristi predviđena vjerojatnost izlaza za taj primjer. Kod višeklasne klasifikacije (eng. *Multi-Class Classification*) računa se gubitak unakrsnom entropijom između očekivane i predviđene distribucije vjerojatnosti klasa. Kod očekivane distribucije se zapravo radi o vektoru čiji jedan element ima vrijednost 1, dok ostali imaju 0. Takav vektor nazivamo *one-hot* vektor. Distribuciju predviđenih vjerojatnosti klasa najčešće radi funkcija softmax. Softmax je aktivacijska funkcija koja normalizira vjerojatnosnu razdiobu klasa, odnosno određuje vjerojatnosti svih klasa na način da njihov zbroj bude 1. Najčešće se koristi na samom izlazu iz modela, a naziv je dobila jer, osim normalizacije, naglašava veće vrijednosti, a smanjuje manje vrijednosti. Ako je ukupan broj klasa  $K$ , funkcija softmax za svaku klasu glasi:

$$f_s(z_k) = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}} \quad (3.6)$$

gdje je  $z$  jedan od izlaza iz modela (vrijednost za jednu od klasa), kolokvijalno zvani *logit*. Funkcija softmax, dakle, prima *logit*, a kao rezultat daje vjerojatnost da se klasa tog *logita* nalazi u ulaznom primjeru. Unakrsnu entropiju za višeklasnu klasifikaciju tada možemo zapisati kao:

$$L = - \sum_{j=1}^K y^{(j)} \log(f_s(z_k)). \quad (3.7)$$

Budući da se kod očekivanog izlaza radi o *one-hot* vektoru pa je vrijednost  $y^{(i)}$  u svim slučajevima osim jednom 0, izraz se može pojednostaviti:

$$L = - \log(f_s(z_T)) \quad (3.8)$$

gdje je  $z_T$  izlaz mreže za očekivanu klasu.

Kad ulazni primjeri mogu pripadati više klasa, koristi se višeznačna klasifikacija (eng. *Multi-Label Classification*). Ukupni gubitak u tom slučaju računa se kao aritmetička sredina gubitaka binarnom unakrsnom entropijom za svaku klasu. Gubitak binarnom entropijom za klasu  $j$  zapisujemo na sljedeći način:

$$L_{(j)} = -y \log(P(Y = 1|X)) - (1 - y) \log(1 - P(Y = 0|X)). \quad (3.9)$$

U tom izrazu  $y$  predstavlja očekivani izlaz klase,  $P(Y = 1|X)$  vjerojatnost dobivenu primjenom funkcije softmax da je izlaz modela za tu klasu 1, a  $P(Y = 0|X)$  da je izlaz modela za tu klasu 0. Ukupni gubitak se može zapisati:

$$L = \frac{1}{K} \sum_{j=1}^K L_{(j)} \quad (3.10)$$

### Aktivacijske funkcije

Aktivacijske funkcije unose nelinearnost u modele kako bi bolje modelirali nelinearna ponašanja promatranih podataka. Osim prije spomenute funkcije softmax koja se nalazi na izlazu modela, u skrivenim slojevima modela dubokog učenja često se koristi zglobnica (eng. *Rectified Linear Unit*, ReLU). Ta aktivacijska funkcija za pozitivni ulaz vraća istu vrijednost, a za negativni 0, odnosno može se zapisati kao:

$$g(z) = \max\{0, z\} \quad (3.11)$$

gdje je  $z$  ulaz u aktivacijsku funkciju. Grafički je funkcija ReLU prikazana na slici 3.5.

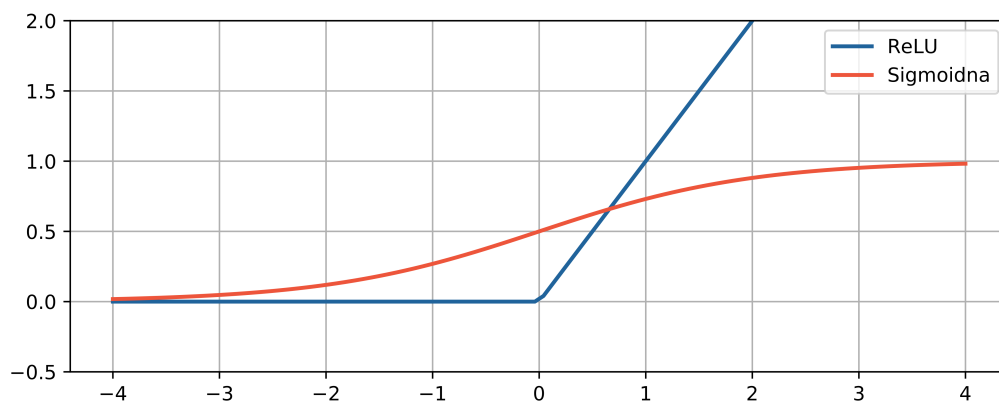
Budući da pozitivne vrijednosti ulaza na izlazu daju istu vrijednost, odnosno funkcija djelom ima linearnu formu, korištenje aktivacijske funkcije ReLU rješava problem nestajućeg gradijenta. Problem nestajućeg gradijenta opisuje smanjenje utjecaja prvih slojeva mreže na izlaz modela do kojeg dolazi zbog prije spomenutog algoritma s povratnim postupkom. U tom algoritmu ažuriranje težina nizom parcijalnih derivacija ovisi i o gradijentu aktivacijske funkcije što u dubljim neuronskim mrežama može smanjiti intenzitet promjene težina prvih slojeva, odnosno može dovesti do situacije da se te težine gotovo uopće ne mijenjaju. Kao jedno od rješenja pokazalo se korištenje ReLU aktivacijske funkcije čija je derivacija 1 za pozitivne vrijednosti ulaza i 0 za negativne što ublažava "nestajanje" gradijenta.

Na izlazu modela s dvije klase često se kao aktivacijska funkcija koristi sigmoidna (logistička) funkcija:

$$f_l(z_k) = \frac{1}{1 + e^{-z_k}}. \quad (3.12)$$

Ta funkcija imitira funkciju praga zbog čega ima svojstvo da primjerima jedne klase daje izlaz

'1', a primjerima druge izlaz '0'. Ono što ju razlikuje od funkcije praga je činjenica da je glatka što znači da je derivabilna na cijeloj domeni, a to je važno za optimizaciju. Grafički prikaz sigmoidne funkcije dan je na slici 3.5.



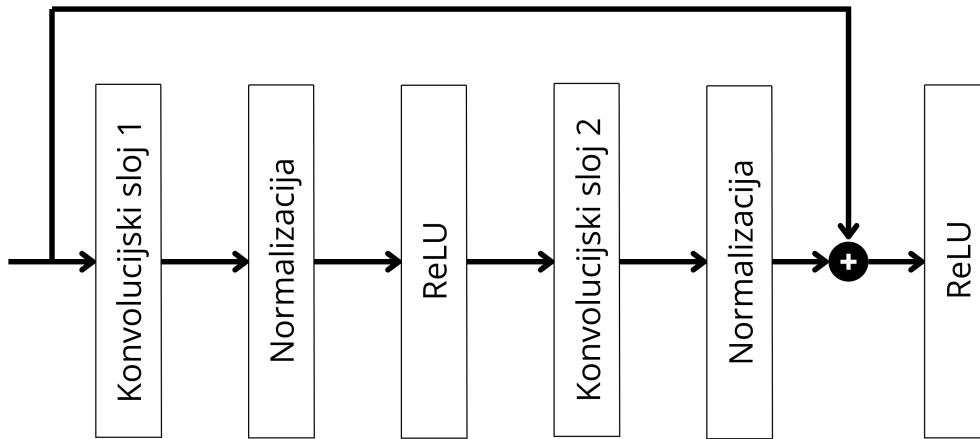
Slika 3.5: Aktivacijske funkcije ReLU (plava) i Sigmoidna (crvena).

## 3.2 Klasifikacijski modeli

U ovom dijelu opisane su najvažnije karakteristike modela koji su korišteni za klasifikaciju objekata na sirovim radarskim podacima i rekonstruiranim slikama. Naglasak je stavljen na ResNet-18 na kojem se temelji većina modela jer je u inicijalnoj usporedbi s ostalim arhitekturama ostvario najbolje rezultate. Modeli su odabrani zbog dobrih rezultata u sličnim primjenama, ali i svoje efikasnosti u implementaciji na uređajima manje procesorske snage.

### 3.2.1 ResNet-18

Suprotno očekivanom, dodavanje slojeva dubokim neuronskim mrežama nakon nekog vremena počinje negativno utjecati na rezultat klasifikacije [87, 88]. Uz problem nestajućeg gradijenta, treniranje jako dubokih mreža procesorski je i vremenski zahtjevno. Potencijalnu arhitekturu takvih mreža koje brže uče, a čija se greška treniranja ne povećava dodavanjem novih slojeva predstavio je Kaiming He na natjecanju *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) 2015. i kasnije s koautorima u radu „Deep Residual Learning for Image Recognition“ [3] koji je do danas citiran 200 000 puta. Rad uvodi ideju rezidualne veze (eng. *Residual Connection*) zbog koje je cijela arhitektura nazvana Rezidualna mreža (eng. *Residual Network*, ResNet), a koja omogućava zaobilažnje slojeva. Autori su krenuli od teze kako bi dublje mreže (s više slojeva) trebale imati bar jednaku pogrešku treniranja kao i pliće, a nikako ne veću. Zato su modelu omogućili prijenos jednake informacije između slojeva putem rezidualne veze. Jedan rezidualni blok prikazan je na slici 3.6.



**Slika 3.6:** Rezidualni blok arhitekture ResNet.

Rezidualni blok u arhitekturi ResNet sastoji se od dva konvolucijska sloja povezana na način da je osim uobičajenog ponašanja u kojem se izlaz prvog sloja nakon ReLU aktivacijske funkcije dovodi na ulaz drugog, kod izlaza drugog sloja se prije provođenja aktivacijske funkcije dovodi i ulaz prvog sloja, odnosno jednom se vezom preskaču ta dva sloja. Za sloj čiji je ulaz  $a[l]$ , izlaz nakon konvolucije i aktivacijske funkcije, odnosno ulaz sljedećeg sloja bit će:

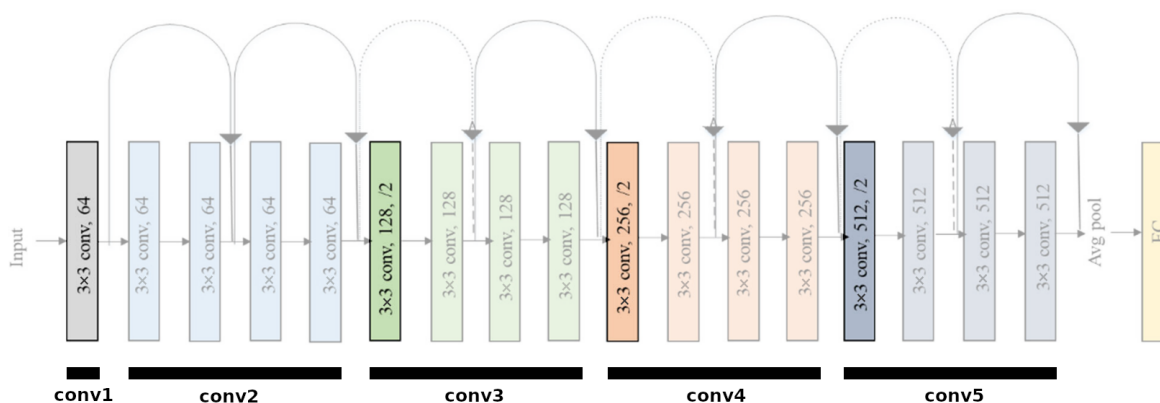
$$a^{[l+1]} = g(z^{[l+1]}) \quad (3.13)$$

gdje je  $z^{[l+1]} = w^{[l+1]}a^{[l]} + b^{[l]}$ , a  $g(x)$  aktivacijska funkcija ReLU. Na sličan način računamo i izlaz drugog sloja rezidualnog bloka, ali mu u aktivacijsku funkciju dodajemo i ulaz prvog sloja:

$$a^{[l+2]} = g(z^{[l+2]} + a^{[l]}). \quad (3.14)$$

Na taj način ako  $z^{[l+2]} \rightarrow 0$ , izlaz rezidualnog bloka može biti njegov ulaz  $a^{[l]}$  jer je, prema jednadžbi 3.11  $g(a^{[l]}) = a^{[l]}$  za pozitivne vrijednosti. Ako se poveže nekoliko rezidualnih blokova, tada se taj ulaz može proslijediti i dublje u mrežu. Svaki konvolucijski sloj uparen je s jednim slojem normalizacije nad grupom (eng. Batch Normalization) [89] koji normalizira ulaz u aktivacijsku funkciju kako bi se ubrzao proces učenja. Treba naglasiti da se u inačicama ResNet mreža koje imaju više slojeva (ResNet-50, ResNet-101, ResNet-152) rezidualni blok sastoji od tri konvolucijska sloja. U implementacijama modela opisanih u ovom radu koristio se ResNet-18 osim kad nije drugačije napomenuto. ResNet-18 ima 17 slojeva podijeljenih u 5 grupa: *conv1*, *conv2*, *conv3*, *conv4*, *conv5* i jedan potpuno povezan sloj. U grupi *conv1* radi se konvolucija filtrom dimenzija 7x7 i sažimanje izborom najvećeg elementa. Sljedeće se četiri grupe sastoje od po dva rezidualna bloka, odnosno u svakoj su grupi četiri konvolucijska sloja s dimenzijama filtra 3x3. U grupama *conv3*, *conv4* i *conv5* podaci se poduzorkuju prvim ko-

nvolucijskim slojem korakom poduzorkovanja 2 čime se dimenzije matrice značajki prepolove, dok se dubina dvostruko povećava. Konvolucijski slojevi grupe *conv2* ne poduzorkuju ulaznu matricu jer su njene dimenzije već smanjene slojem sažimanja koji se nalazi na kraju grupe *conv1*. Izlaz posljednje grupe (*conv5*) sažima se prema prosječnoj vrijednosti elemenata ravna i prosljeđuje potpuno povezanom sloju koji s aktivacijskom funkcijom softmax daje izlaz modela. Slika 3.7 prikazuje cijelu arhitekturu ResNet-18 mreže s naglašenim slojevima u kojima se odvija poduzorkovanje.



Slika 3.7: Arhitektura mreže ResNet-18 [80].

### 3.2.2 LSTM i MobileNetV3

Osim ResNeta, za usporedbu rezultata modela različitih arhitektura prilagođenih radu na računalima slabijih karakteristika u koje se može uvrstiti i Raspberry Pi, korištene su LSTM i MobileNetV3 mreže.

LSTM (od eng. *Long Short-Term Memory*) vrsta je povratnih neuronskih mreža (eng. *Recurrent Neural Network*, RNN) koje rade sa slijednim podacima na način da osim ulaza u skrivenom sloju primaju i izlaz prošlog primjera. Takve se mreže često koriste u obradi jezika u što spadaju obrada teksta i govora. Iz perspektive ovog rada posebno je zanimljiva primjena LSTM-a u obradi govora u stvarom vremenu jer su ulazni podaci u obliku slijednih signala što je slično rezultantskim signalima dobivenim na svakoj poziciji senzora GBSAR sustava. Ti se signali pojedinačno dovode na ulaz LSTM mreže gdje se prvo u *Embedding* sloju preoblikuju u reprezentacijski vektor prilagođen za LSTM sloj. *Embedding* sloj se u obradi teksta koristi za enkodiranje ulaznih podataka, odnosno riječi, u njihovu reprezentaciju vektorima, a u našem će slučaju signale prikazane u obliku vektora s 1024 elementa pretvarati u vektore s 256 elemenata. U LSTM neuronskoj mreži taj sloj uči reprezentacije ulaznih podataka i prilagođava ih zadatku koji mreža rješava, dok LSTM sloj prema naučenim reprezentacijama određuje koje njihove dijelove mreža treba zapamtiti, a koje može zaboraviti jer ne doprinose konačnom rezultatu. Nakon LSTM sloja je, slično kao u CNN-u, potpuno povezani sloj.

MobileNetV3 je vrsta CNN-a optimizirana za procesorske jedinice mobilnih uređaja. Arhitekturu te mreže odredio je Googleov *Neural Architecture Search* (NAS) [90], postupak za automatsko pronalaženje optimalnih arhitektura neuronskih mreža temeljem postavljenih zahtjeva. Za MobileNetV3 zahtjev NAS-u je, osim preciznosti, bila i efikasnost izvođenja na mobilnom hardveru. Osnovna arhitektura dodatno je optimizirana algoritmom NetAdapt [91] koji ju iterativno podrezuje (eng. *Prunning*) smanjujući broj izlaznih kanala nekog konvolucijskog sloja, korištenjem aktivacijske funkcije *HardSwish* umjesto ReLU i bloka „*squeeze and excitation*“ koji uči dinamički rekalibrirati izlazne kanale konvolucijskih slojeva. MobileNetV3 mreža korištena je za klasifikaciju objekata u oba pristupa – i sirovim radarskim podacima i rekonstruiranim slikama, dok je LSTM mreža, budući da je prilagođena slijednim podacima, testirana samo sa sirovim rezultantskim signalima GBSAR-a.

### 3.3 Metrike

U višeznačnoj klasifikaciji očekivano i predviđeno stanje svake klase može biti '0' ili '1' gdje '0' predstavlja da se objekt ne nalazi na primjeru, a '1' da se nalazi. To znači da se za evaluaciju modela mogu koristiti metrike koje rade s binarnim vrijednostima. Predviđanje za neku klasu može se svrstati u jednu od četiri grupe:

- točno negativni: TN (od eng. *True Negative*) - model je ispravno klasificirao da objekt nije prisutan na primjeru,
- pogrešno negativni: FN (od eng. *False Negative*) - model je pogrešno klasificirao da objekt nije prisutan na primjeru, a jest,
- točno pozitivni: TP (od eng. *True Positive*) - model je ispravno klasificirao da je objekt prisutan na primjeru,
- pogrešno pozitivni: FP (od eng. *False Positive*) - model je pogrešno klasificirao da je objekt prisutan na primjeru, a nije.

Pomoću tih grupa može se izračunati nekoliko metrika koje na različite načine evaluiraju model, a u nastavku će se ukratko opisati one koje se koriste u radu:

- Točnost (eng. *Accuracy*) prikazuje omjer točno predviđenih primjera u odnosu na sve. To je vjerojatno najintuitivnija mjera za evaluaciju, ali se pokazalo da daje preoptimistične rezultate kod skupova koji nisu uravnoteženi, odnosno koji imaju različitu distribuciju primjera po klasama [92]. Točnost računamo prema sljedećem izrazu:

$$Acc = \frac{TN + TP}{TN + FN + TP + FP}. \quad (3.15)$$

- Preciznost (eng. *Precision*) prikazuje omjer broja ispravno klasificiranih u kojima se objekt nalazio na primjerima u odnosu na sve za koje je model predvidio da se nalaze na



primjerima (svi pozitivni). Preciznost se smanjuje kad se broj pogrešno pozitivnih (FP) povećá:

$$P = \frac{TP}{TP + FP}. \quad (3.16)$$

- Odziv (eng. *Recall*) prikazuje omjer broja ispravno klasificiranih u kojima se objekt nalazio na primjerima u odnosu na sve koji su se zapravo nalazili na primjeru, odnosno čije je očekivanje 1. Odziv se smanjuje kad se broj pogrešno negativnih (FN) povećá:

$$R = \frac{TP}{TP + FN}. \quad (3.17)$$

- F1 mjera (eng. *F1-score*) pojednostavljeni je oblik F mjere (eng. *F-score*) [92] u kojem je  $\beta = 1$  i prikazuje harmonijsku sredinu preciznosti i odziva. F1 mjera se dodatno smanjuje ("kažnjava") pogrešno predviđenim klasifikacijama:

$$F1 = \frac{2 \cdot P \cdot R}{P + R} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}. \quad (3.18)$$

- Makro F1 mjera (eng. *macro F1-score*) koristi se kod višeklasne klasifikacije, a računa se kao aritmetička sredina F1 mjere za svaku klasu:

$$mF1 = \frac{1}{K} \sum_{j=1}^K F1_k. \quad (3.19)$$

Spomenuto je ranije kako je izlaz modela vjerojatnost da se objekt nalazi na primjeru. Postavlja se pitanje kolika vjerojatnost u višeo značajnoj klasifikaciji treba biti da se za određeni objekt predvidi da se on zaista i nalazi na tom primjeru. Intuitivno je logično koristiti klasifikacijski prag (eng. *threshold*) od 0.5 što znači da će model klasu čija je vjerojatnost viša od 0.5 označiti kao '1', a klasu vjerojatnosti ispod 0.5 kao '0'. Ipak, pokazalo se da to nije najbolje rješenje za sve primjene. S druge strane, povećanje praga povećat će preciznost jer će samo primjeri s visokom vjerojatnošću biti klasificirani kao '1', a snižavanje će povećati odziv jer će u većem broju slučajeva primjere klasificirati kao pozitivne što će rezultirati manjim brojem lažno negativnih. U praksi su preciznost i odziv često suprostavljeni na način da visoka preciznost za sobom povlači niži odziv, a, očekivano, cilj svakog modela je što je moguće više približiti se rezultatu  $P = 1$  i  $R = 1$ . Zato se klasifikacijski prag postavlja pomoću krivulje preciznost-odziv. Takva se krivulja, kojoj su na osi  $y$  preciznosti, a na osi  $x$  odzivi, generira postavljanjem različitih vrijednosti praga za koji se potom računaju preciznost i odziv, a integral krivulje daje novu metriku zvanu prosječna preciznost (eng. *Average Precision*, AP) [93]. U višeklasnoj klasifikaciji kao konačni rezultat evaluacije modela uzima se aritmetička sredina prosječnih preciznosti svih klasa. Ta se metrika naziva srednja AP vrijednost (eng. *mean Average Precision*, mAP).

## Poglavlje 4

### Kreirani skupovi podataka

Izrađeno je i javno objavljeno pet skupova radarskih podataka snimljenih razvijenim radarom sa sintetičkom aperturom koji se međusobno razlikuju u radarskim parametrima korištenim tijekom snimanja i promatranim scenama. U prvom su objedinjeni parovi sirovih radarskim podataka i rekonstruiranih slika, u drugom su snimljene scene u kojima su promatrani objekti postavljeni iza prepreke, u trećem su iste scene snimljene korištenjem različitih radarskih parametara, u četvrtom je promatrano devet različitih objekata, a u petom su iste scene snimane različitim polarizacijama. Objekti su u svim skupovima postavljeni na udaljenost manju od 1 m od radara. Svi su skupovi javno dostupni na mrežnoj stranici Mendeley Data [61, 62, 63, 64, 65].

Kreiran je još jedan skup podataka nazvan LabSAR koji nije dodatno opisan. Mjerenja tog skupa napravljena su u kontroliranim uvjetima u neječnoj komori, a u scenama su se nalazili testni objekti koji su se jako razlikovali što je u gotovo svim situacijama davalo nerealno visoku točnost klasifikacije. Ipak, zbog sličnih svojstava kao i kasniji skupovi podataka, LabSAR je korišten za postavljanje početnih težina modela.

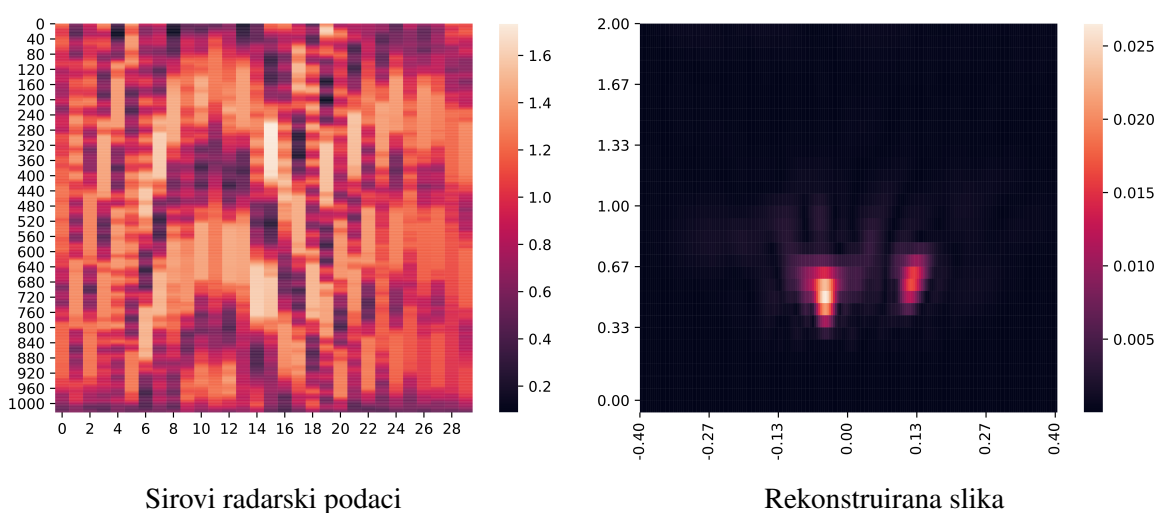
#### 4.1 Parovi sirovih radarskih podataka i rekonstruiranih slika

Repozitorij se sastoji od dva dijela: RealSAR-RAW i RealSAR-IMG. U prvom se nalaze sirovi podaci dobiveni radarom u formatu tekstualne datoteke, a u drugom slike u formatu PNG (eng. *Portable Network Graphics*) rekonstruirane primjenom Omega-K algoritma nad tim podacima. U provedenim mjerenjima pomoću razvijenog radara GBSAR-Pi pomicala je platforma sa senzorom 30 cm u načinu rada stani-kreni (eng. *stop-and-go*) s duljinom koraka od 1 cm pritom odašiljući chirp signale centralne frekvencije 24 GHz i širine pojasa 700 MHz predstavljenih pomoću 1024 frekvencijska uzorka. Prema tome, rezultatna matrica dobivena nakon snimanja jedne scene je dimenzija 1024x30. U mjerenjima je nasumično korištena horizontalna i vertikalna polarizacija. Kako bi se povećao odnos signal-šum, u svakom je GBSAR koraku (na svakoj poziciji senzora) ukupno emitirano 10 chirpova čiji se rezultat nakon miješanja s

primljenim signalom usrednjio.

U promatranoj sceni nalazila su se tri testna objekta sličnih dimenzija, ali različitih materijalnih svojstava: jedan od plastike, drugi od stakla, a treći od metala. Kako bi se otežao taj zadatak, scene su namjerno snimane u prostoriji s raznim elementima koji su reflektiranim signalima dodali šum, a i sami su objekti postavljeni na razne pozicije unutar radarski promatranog područja. Scena se mogla sastojati od jedne od osam kombinacija s tri objekta. Tablica 4.1 prikazuje informacije o radaru i snimljenim scenama.

Ukupno je snimljeno 337 scena, odnosno skup ima 337 parova sirovih radarskih podataka i rekonstruiranih slika. U 172 scene je bio postavljen metalni objekt, u 172 stakleni, u 179 plastični, a 29 scena je snimljeno bez objekta. Primjer jednog para u kojem su se u promatranoj sceni nalazili metalni i stakleni objekt dan je na slici 4.1. Repozitorij je javno dostupan na [64].



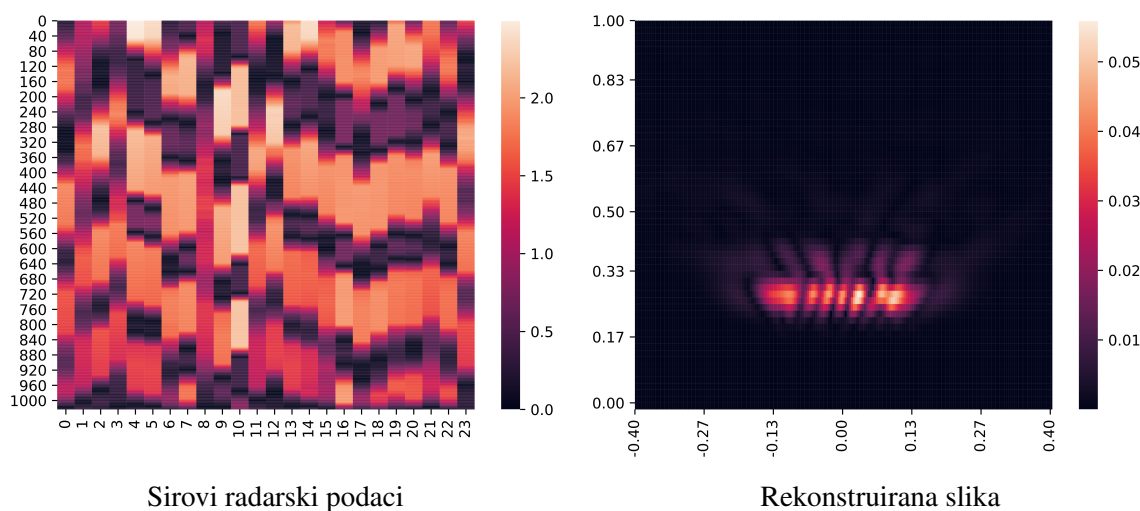
**Slika 4.1:** Primjer jednog para sirovih radarskih podataka u obliku matrice i rekonstruirane SAR slike prikazanih toplinskom mapom. U promatranoj sceni snimljeni su metalni i stakleni objekt. Primjetna je razlika u intenzitetu refleksije.

## 4.2 Scene snimljene iza prepreke

Repozitorij sadrži sirove radarske podatke (snimljene pomoću razvijenog radara) scena u kojima su objekti postavljeni iza kartonske pregrade. Svaka je scena snimljena u 24 koraka što znači da je rezultatna matrica koja služi kao ulaz u model za klasifikaciju dimenzija 1024x24. Centralna frekvencija FMCW radara je 24 GHz, a širina pojasa 1.3 GHz. Sirovi podaci u obliku matrice i rekonstruirana slika prikazani su na slici 4.2

Ideja iza kreiranja ovog skupa podataka bila je analizirati mogućnost detekcije objekata s defektom u paketu pa su objekti podijeljeni u dvije klase – jedna koja predstavlja objekt s defektom i druga koja predstavlja normalan objekt. U svakoj je sceni ukupno bilo 5 objekata

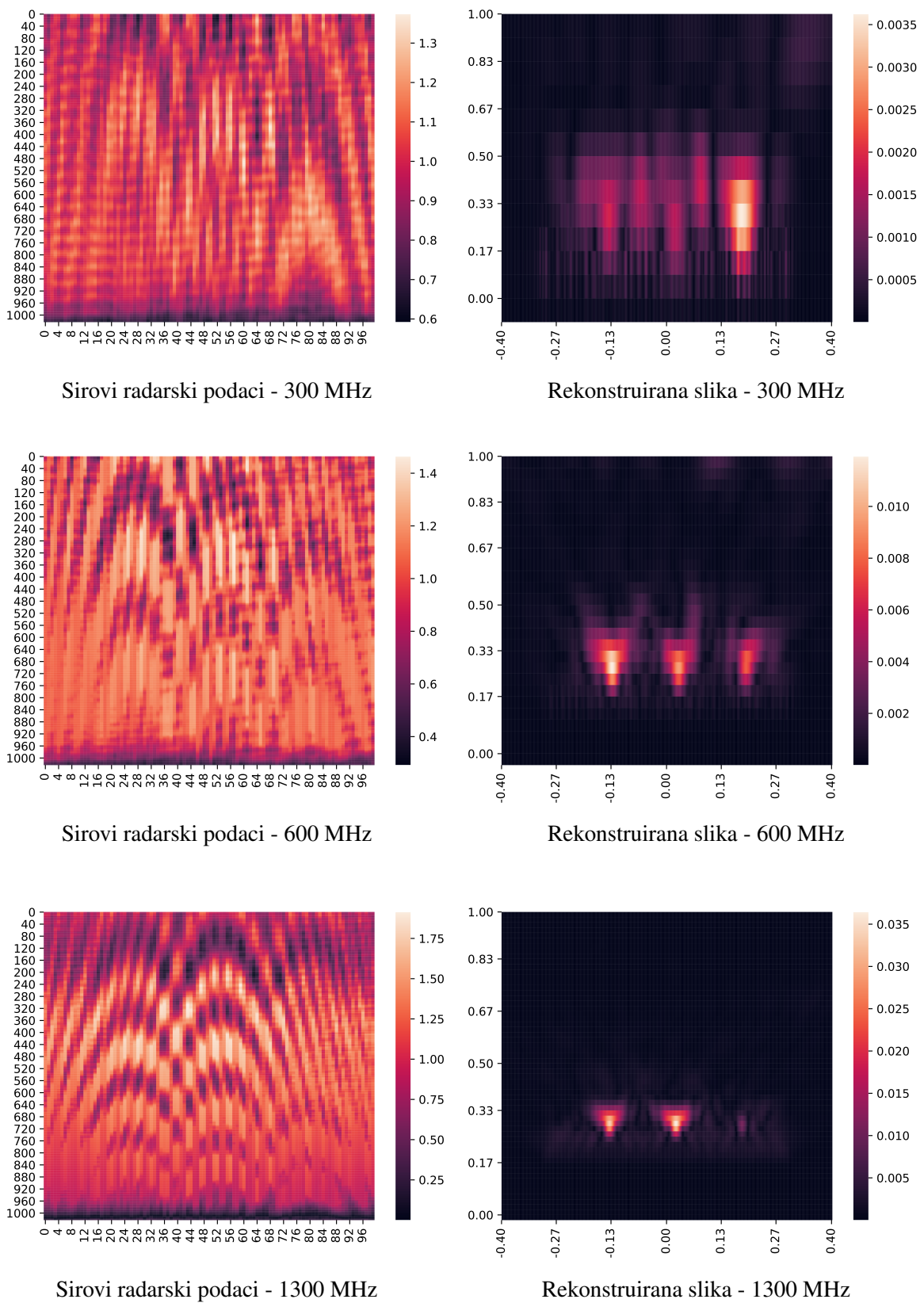
od kojih je jedan potencijalno bio onaj s defektom. Od 155 snimljenih scena, 50 ih je bilo bez takvog objekta, odnosno 105 scena je uz četiri normalna, imalo jedan objekt s defektom. Skup podataka osim za detekciju može se koristiti i za lokalizaciju objekta budući da je u scenama s defektnim objektom označena i njegova pozicija. Repozitorij je javno dostupan na [63].



**Slika 4.2:** Primjer sirovih radarskih podataka (lijevo) scene u kojoj su metalni objekti postavljeni iza prepreke. Na rekonstruiranoj slici (desno) vide se objekti, ali je njihovo raspoznavanje otežano. U konkretnoj sceni defektni objekt nalazio se na srednjoj poziciji.

### 4.3 Scene snimljene radarom s različitim parametrima

Repozitorij sadrži sirove radarske podatke snimljene pomoću razvijenog radara kojem su postavljene različite kombinacije parametara duljine koraka i širine frekventijskog pojasa. Ti su parametri odabrani kako bi se na podatke utjecalo u dva prostorna smjera – po azimutu i po udaljenosti. Ukupan put koji je platforma sa senzorom prelazila tijekom mjerenja nije se mijenjao, ali broj koraka, odnosno udaljenost između dva uzastopna mjerenja jest. Tako je put duljine 50 cm prijeđen u 100, 50, 33, 17 i 14 koraka što predstavlja duljinu koraka od 0.5, 1, 1.5, 3 i 3.5 cm. S druge strane, u svakom je koraku FMCW radar (senzor) radio mjerenje koristeći pritom jednu od tri širine pojasa – 300, 600 i 1300 MHz. Odnosno, u svakom koraku poslano je 10 chirp signala širine pojasa od 300 MHz, zatim 10 širine 600 MHz i na kraju 10 širine pojasa 1300 MHz. Za svaku se širinu pojasa spremala usrednjena vrijednost rezultatnog signala u tom koraku. Na kraju su pohranjene matrice dimenzija 1024xbroj koraka.



**Slika 4.3:** Primjeri sirovih radarskih podataka i rekonstruiranih slika za tri širine pojasa (300, 600 i 1300 MHz). U promatranoj sceni gumeni objekt nalazio se u trećem keramičkom objektu (klasa 7).

U promatranim scenama postavljena su tri keramička objekta unutar kojih se nalazio metalni objekt, gumeni objekt ili ništa. U nijednoj se sceni nisu nalazili i metalni i gumeni objekt istovremeno pa je ukupni broj klasa 7:

1. Tri prazna keramička objekta
2. Metalni objekt u prvom keramičkom objektu
3. Metalni objekt u drugom keramičkom objektu
4. Metalni objekt u trećem keramičkom objektu
5. Gumeni objekt u prvom keramičkom objektu
6. Gumeni objekt u drugom keramičkom objektu
7. Gumeni objekt u trećem keramičkom objektu

Ideja iza kreiranja ovakve baze radarskih podataka ponovno je klasifikacija objekata skrivenih iza drugih, u ovom slučaju keramičkih. Od 243 originalnih mjerenja u kojima je korištena duljina koraka od 0.5 cm (81 po širini frekvencijskog pojasa), umjetno su, uklanjanjem mjerenja s određenih pozicija, generirane matrice za ostale duljine koraka. Baza se sastoji od ukupno 10935 sirovih radarskih podataka i javno je dostupna na [62]. Sirovi podaci i rekonstruirane slike prikazane su na slici 4.3. Na tom je primjeru promatrana scena broj 7 (gumeni objekt u trećem keramičkom). Osim lošije rezolucije kod 300 i 600 MHz, može se primijetiti kako se na slici rekonstruiranoj podacima snimljenim korištenjem najveće širine frekvencijskog pojasa (1300 MHz) treći objekt gotovo i ne vidi što implicira da pri toj širini pojasa potencijalno dolazi do izražene interferencije i poništavanja signala.

## 4.4 Scene s 9 različitih objekata

Skup podataka sadrži 267 radarskih mjerenja. Promatrane su scene s tri objekta različitih dimenzija i materijala. Ukupno je korišteno 9 objekata – 4 metalna, 3 staklena i 2 plastična. Scena je snimljena sa 120 koraka kojima je pokriveno 60 cm, odnosno duljina koraka je 0.5 cm. Dobljena matrica je nakon toga podijeljena u tri jednaka dijela kako bi svaki objekt bio u zasebnoj matrici dimenzija 1024x40.

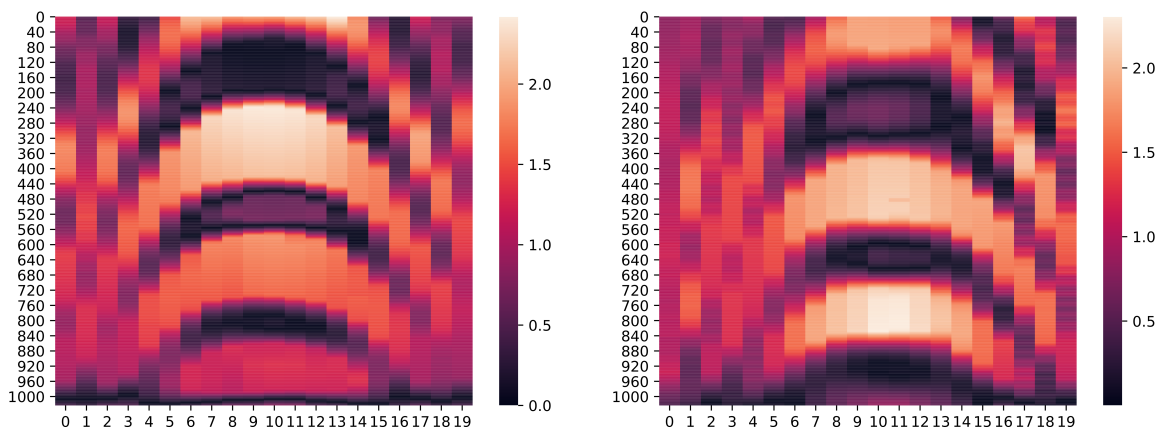
Polarizacija FMCW radara je vertikalna, centralna frekvencija 24 GHz, a širina pojasa 1.3 MHz. Repozitorij je javno dostupan na [61].

## 4.5 Scene snimljene različitim polarizacijama

Skup sadrži 150 parova sirovih radarskih podataka snimljenih horizontalnom i vertikalnom polarizacijom. Slično kao u prethodnom skupu, promatrane su scene imale tri objekta, a rezultantske su matrice potom podijeljene u tri jednaka dijela. U ovom su slučaju scene promatrane sa 60 pozicija međusobno udaljenih 1 cm, što znači da je originalna matrica dimenzija 1024x60,

odnosno da su dijelovi matrice koji sadrže po jedan objekt dimenzija 1024x20.

Promatrani su objekti različitog materijala (metal, staklo, plastika) i oblika (valjak i kvadar). Sve su scene prvo snimane horizontalnom pa onda vertikalnom polarizacijom tako da je ukupan broj mjerenja 300 (150 po polarizaciji). Toplinske mape jednog para ovog skupa i njihove rekonstruirane slike prikazane su na slici 4.4. Repozitorij je javno dostupan na [65].



Sirovi podaci - horizontalna polarizacija

Sirovi podaci - vertikalna polarizacija

**Slika 4.4:** Primjeri sirovih radarskih podataka snimljenih horizontalnom (lijevo) i vertikalnom (desno) polarizacijom. U tom dijelu scene nalazio se metalni kvadar [94].

**Tablica 4.1:** Pregled parametara mjerenja radarom GBSAR-Pi kojima su kreirani skupovi podataka

Skup	Scena	Broj mjerenja	Širina pojasa	Duljina koraka	Ukupna apertura
4.1	3 objekta različitih materijala	337	700 MHz	1 cm	30 cm
4.2	5 objekata, 1 s defektom, postavljeni iza prepreke	155	1300 MHz	1 cm	24 cm
4.3	2 objekta različitih materijala postavljeni unutar keramičkih	10935	300 MHz 600 MHz 1300 MHz	0.5, 1, 1.5 3, 3.5 cm	50 cm
4.4	9 objekata različitih materijala	267	1300 MHz	0.5 cm	60 cm
4.5	6 objekata različitih materijala	300	1300 MHz	1 cm	60 cm

## Poglavlje 5

# Klasifikacija objekata pomoću sustava

## GBSAR-Pi

Analizirani su rezultati klasifikacijskih modela trenirani i testirani na prije navedenim skupovima. Prvi skup podataka omogućio je usporedbu dvaju pristupa klasifikaciji – prvi je temeljen na sirovim podacima, a drugi na rekonstruiranim slikama. Skupom radarskih podataka snimljenih s različitim radarskim parametrima uspoređen je utjecaj promatranih parametara na točnost klasifikacije. Podacima četvrtog skupa testirana je ideja povratne veze između klasifikacijskog modela i radara kojom je planirano povećati energetske učinkovitost sustava, a petim mogućnost dobivanja dodatne informacije za model korištenjem različitih polarizacija prilikom snimanja.

### 5.1 Usporedba pristupa sirovim podacima i rekonstruiranim slikama

Algoritam za rekonstrukciju slike opisan u poglavlju 2.3.2 zbog aproksimacija i Stoltove interpolacije narušava integritet sirovih podataka i unosi određenu nepreciznost koja utječe na točnost klasifikacije. Kako bi se testiralo do koje mjere Omega-K algoritam utječe na integritet podataka iz perspektive klasifikacije, uspoređeni su rezultati dvaju modela: prvi je temeljen na sirovim radarskim podacima, a drugi na slikama rekonstruiranim pomoću tih istih radarskih podataka. Odnosno, korišten je skup podataka predstavljen u poglavlju 4.1.

Klasifikacija je rađena prema višeznačnoj paradigmi (eng. *Multi-Label*) u kojoj model odjednom raspoznaje sve objekte u sceni. Ideja je da model u prvom dijelu procesa treniranja nauči zajednička svojstva, a kasnije se razgrana u nekoliko neovisnih binarnih klasifikacija. U ovom slučaju takvih je grana koliko i objekata (svaki objekt je jedna klasa), a svaka klasificira prisutnost jednog objekta u sceni. Budući da se radi o višeznačnoj klasifikaciji, gubitak modela računa se kao aritmetička sredina gubitaka binarnih unakrsnih entropija svih grana što je detalj-



nije opisano u poglavlju 3.1.2. U našem slučaju u kojem ima više objekata na jednom ulaznom primjeru takva klasifikacija za svaki objekt ima jednu klasu, odnosno za  $n$  objekata ima  $n$  izlaza iz modela jer se za svaki objekt zasebno određuje nalazi li se na ulaznom primjeru ili ne. S druge strane, u višeklasnoj klasifikaciji svaka moguća kombinacija objekata predstavlja jednu klasu. Takvih je kombinacija za  $n$  objekata  $2^n$  pa je istraživanje nastavljeno primarno u smjeru višeo značne klasifikacije. Budući da je skup podataka na kojem su testirana oba pristupa kreiran s tri objekta, izlaz iz oba pristupa (sirovi podaci i rekonstruirane slike) jesu tri vjerojatnosti za prisutnost određenog objekata u sceni.

Tijekom procesa treniranja umjetno se povećala varijabilnost ulaznog skupa (augmentacija) kako bi mreža imala više primjera za učenje i spriječila pretreniranost. Augmentacija podataka u našem se slučaju radila okretanjem slike i nasumičnom promjenom vrijednosti RGB kanala u pristupu rekonstruiranim slikama, odnosno dodavanjem Gaussovog šuma s varijancom  $\sigma^2 = 0.01$  svakom elementu matrice. U svim je eksperimentima korišten optimizacijski postupak Adam [11], stopa učenja  $2 \cdot 10^{-4}$  i hiperparametar opadanja težina (eng. weight decay)  $3 \cdot 10^{-3}$ . Svaki je model treniran kroz 30 epoha.

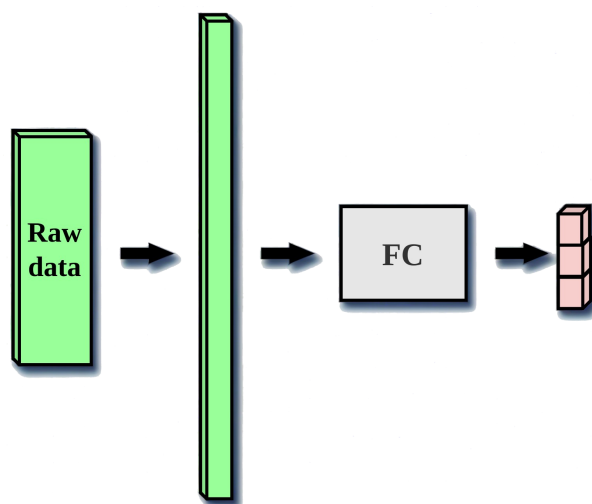
### 5.1.1 Klasifikacijski modeli temeljeni na sirovim radarskim podacima

Za pristup temeljen na sirovim radarskim podacima uspoređene su četiri arhitekture: prva se sastoji od jednog potpuno povezanog sloja, druga koristi LSTM mrežu za procesiranje podataka prije klasifikacije, a treća i četvrta su konvolucijske mreže MobileNetV3 i ResNet-18. Sve navedene arhitekture često se koriste u ugradbenim računalima zbog svoje jednostavnosti i efikasnosti.

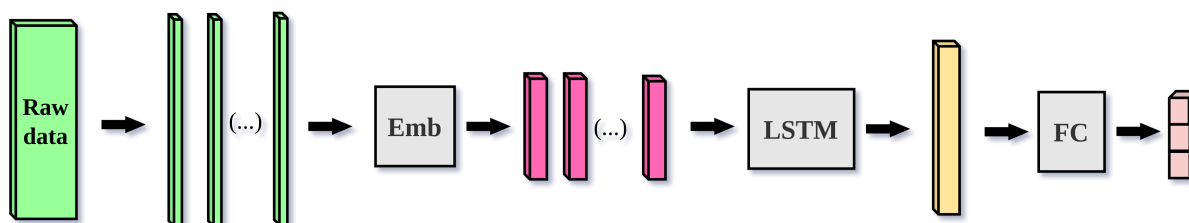
Kod prve arhitekture s jednim potpuno povezanim slojem se matrica dimenzija 1024x30 ravna u vektor od 30 720 elemenata koji se predaje potpuno povezanom sloju, a koji generira izlazni vektor od tri elementa (vjerojatnosti za tri klase) pomoću aktivacijske funkcije sigmoid. Arhitektura te mreže prikazana je na slici 5.1.

U drugoj se arhitekturi koristi dvosmjerna LSTM mreža koja ulaznu matricu (1024x30) promatra kao 30 vektora veličine 1024 elementa, a svaki je od vektora pomoću naučene matrice (embedding sloja) ugrađen u prostor s 256 dimenzija što se smatra standardnim načinom transformacije ulaza u LSTM mrežu. Arhitektura te mreže prikazana je na slici 5.2.

MobileNetV3 i ResNet-18 modeli, iako složeniji od prošla dva, primjenjivi su na računalima male snage kao što je Raspberry Pi. Njihove su se arhitekture, kao i većina arhitektura CNN-a, pokazale vrlo učinkovitim u zadacima klasifikacije na slikama. Budući da su slike zapravo matrice, očekivano je da će takvi modeli imati dobre rezultate i u slučaju niza sirovih signala prikazanih u obliku matrice. Uz to, kod matrica sirovih podataka su, slično kao na rekonstruiranim slikama, vrijednosti susjednih elemenata povezane u jednom smjeru po prostoru (azimut), a u drugom kao niz frekvencijskih uzoraka (udaljenost) što dodatno opravdava smisao



**Slika 5.1:** Arhitektura mreže s jednim potpuno povezanim slojem. Mreža sirove podatke ravna u jedan vektor koji je potpuno povezan s izlazom za tri klase [80].



**Slika 5.2:** Arhitektura mreže LSTM koja 30 signala matrice ulaza prima kao zasebne vektore. Prije ulaza u LSTM sloj vektori se transformiraju pomoću Embedding sloja (Emb), a prije izlaza prolaze još kroz potpuno povezani sloj (FC) [80].

korištenja takvih modela. Iz perspektive ovog rada, primjena na sirovim podacima omogućuje usporedbu istih arhitektura na različitim ulaznim podacima koji predstavljaju istu scenu. Konkretnije, budući da su obje arhitekture modela korištene i na rekonstruiranim slikama, moguće je izravno usporediti rezultate modela MobileNetV3 treniranog na sirovim podacima s onima modela MobileNetV3 treniranog na slikama kao i modela arhitekture ResNet-18 na sirovim podacima s rezultatima modela ResNet-18 treniranog na rekonstruiranim slikama.

U prvom je koraku testiran utjecaj augmentacijskog postupka dodavanja Gaussovog šuma. U tablici 5.1 prikazane su srednje prosječne preciznosti svakog modela s dodanim šumom i bez. Tablica potvrđuje da dodavanje šuma matricama u procesu treniranja povećava preciznost kod svakog modela. Kao što je i očekivano, pokazuje i da složenije arhitekture daju veću srednju prosječnu preciznost pa su u daljnjoj usporedbi, budući da se one koriste i u pristupu rekonstruiranim slikama, promatrane samo te dvije arhitekture. Između njih je model ResNet-18 postigao bolje rezultate. Uz to, primijećeno je da se preciznost može dodatno povećati prilagodbom same arhitekture ulaznim podacima u obliku matrice čije dimenzije nisu uobičajene. U nastavku je opisana modifikacija arhitekture ResNet-18.

**Tablica 5.1:** Usporedba rezultata modela treniranih na sirovim radarskim podacima. Prikazana je mjera srednje prosječne preciznosti (mAP).

Augmentacija	FC	LSTM	MNV3	RN18
Ne	90.89	92.51	96.07	96.95
Da	90.97	92.87	96.23	97.12

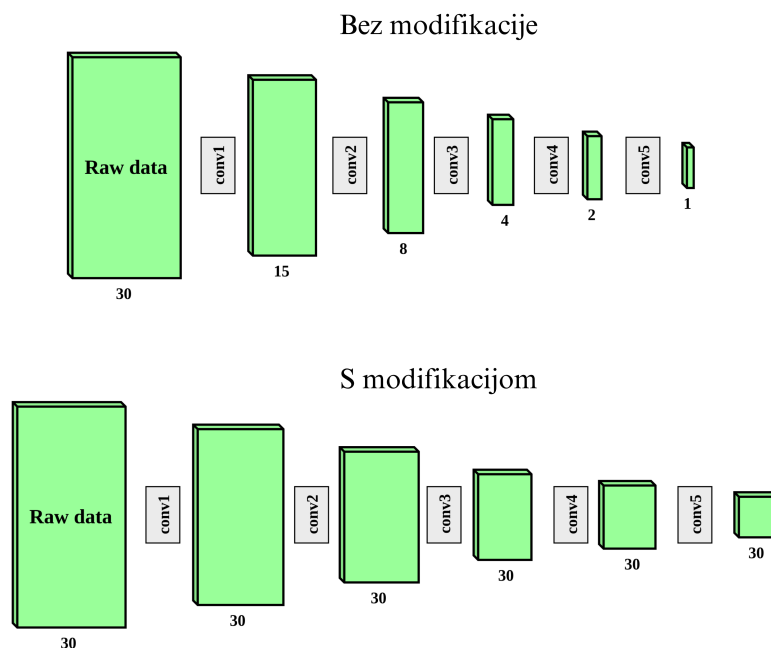
### 5.1.2 Modifikacija arhitekture ResNet-18

Kao što je opisano u poglavlju 3.2.1, ResNet-18 se sastoji od 18 slojeva, a ulazni se podaci poduzorkuju pet puta prije potpuno povezanog sloja:

- sažimanjem izborom najvećeg elementa u grupi *conv1*,
- prvim konvolucijskim slojem u grupama *conv3*, *conv4* i *conv5* i
- sažimanjem prema prosječnoj vrijednosti nakon grupe *conv5*.

Budući da je korak poduzorkovanja 2, to znači da se dimenzije ulazne matrice na svakom od tih pet mjesta upola smanje kako bi se ubrzao proces treniranja. Odnosno dimenzije matrice prije (posljednjeg) potpuno povezanog sloja su 32 puta manje od dimenzija ulazne matrice. Kod rekonstruiranih slika dimenzija 496x369 to smanjenje ne predstavlja problem, ali kod sirovih radarskih podataka takvo smanjenje može izazvati gubitak dijela informacije vezanog uz sličnosti mjerenja FMCW radara u različitim koracima. Konkretnije, kod horizontalne dimenzije matrice koja predstavlja broj koraka GBSAR-a, a koji je u slučaju ovog skupa podataka 30, takvo smanjenje rezultiralo bi matricom (prije potpuno povezanog sloja) čija je horizontalna dimenzija 1 što znači da kasniji slojevi ne bi mogli koristiti potencijalnu informaciju koju daje odnos između dva mjerenja FMCW-a. Iako su rezultati prikazani u tablici 5.1 pokazali da ResNet-18 daje zadovoljavajuću preciznost, odlučili smo testirati kako bi se model ponašao kada bi se prilagodila arhitektura mreže ResNet-18 na način da se ukine poduzorkovanje horizontalne osi matrice tako što se korak poduzorkovanja na prije navedenih pet mjesta postavi na 1. Važno je napomenuti da se korak poduzorkovanja nije mijenjao za vertikalnu dimenziju tako da se ona s 1024 na kraju spustila na 32. Modifikacijom se završna matrica (prije potpuno povezanog sloja) s dimenzija 32x1 promijenila na 32x30 (dubina je ostala 512) što je prikazano na slici 5.3. Nakon provođenja Globalnog prosječnog sažimanja (Global Averaging Pooling) vektor veličine 512 prosljeđuje se potpuno povezanom sloju koji daje izlazni vektor s tri vjerojatnosti za tri klase. Budući da modifikacija ne mijenja dimenzionalnost filtara (kernel) konvolucijskog sloja i ne dodaje nove parametre modelu, predtrenirani set parametara ResNet-18 arhitekture mogu se i dalje koristiti.

U usporedbi s prije navedenim modelima, modificirani ResNet-18 prilagođen za rad sa sirovim radarskim podacima postigao je višu srednju prosječnu preciznost u oba testirana slučaja – s i bez augmentacijskog postupka dodavanja Gaussovog šuma. U tablici 5.2 dana je usporedba



**Slika 5.3:** Modifikacija mreže ResNet-18 na primjeru dimenzija ulazne matrice sa signalima GBSAR-a uklanjanjem koraka poduzorkovanja (eng. downsampling) [80].

**Tablica 5.2:** Usporedba rezultata svih razvijenih modela treniranih na sirovim radarskim podacima. Prikazana je mjera srednje prosječne preciznosti (mAP).

Augmentacija	FC	LSTM	MNV3	RN18	ModRN18
Ne	90.89	92.51	96.07	96.95	<b>99.51</b>
Da	90.97	92.87	96.23	97.12	<b>99.73</b>

svih implementiranih modela temeljenih na sirovim podacima.

### 5.1.3 Klasifikacijski modeli temeljeni na rekonstruiranim slikama

U implementaciji modela treniranih na rekonstruiranim slikama korištene su MobileNetV3 i ResNet-18 arhitekture. Ulaz u oba modela je RGB slika dimenzija (496, 369, 3) dobivena primjenom rekonstrukcijskog algoritma nad sirovim radarskim podacima. Implementacija algoritma opisana je u poglavlju 2.3.2. Budući da se radi o običnim slikama, moguće je koristiti standardne prakse računalnog vida za pripremu podataka za treniranje. Izlaz oba modela dobi-ven je funkcijom sigmoid koja daje vjerojatnosti prisutnosti svakog objekta u sceni.

### 5.1.4 Rezultati usporedbe dvaju pristupa

U pristupu sirovim radarskim podacima uspoređeni su rezultati klasifikacije modela arhitek- ture MobileNetV3, ResNet-18 i modificirani ResNet-18, a u pristupu rekonstruiranim slikama

modeli arhitekture MobileNetV3 i ResNet-18. Kod svih su modela testirana četiri načina postavljanja početnih težina modela:

- Nasumično
- Skupom podataka ImageNet
- Skupom podataka LabSAR
- Skupovima podataka ImageNet i LabSAR

U tablici 5.3 prikazane su srednje prosječne preciznosti svih kombinacija modela i načina postavljanja početnih težina. Rezultati pokazuju da pristup rekonstruiranim slikama (u tablici označene s 'IMG') daje bolje rezultate u odnosu na pristup sirovim podacima (u tablici 'RAW') kad se modeli temelje na standardnim arhitekturama MobileNet i ResNet-18 što je očekivano budući da su te arhitekture prilagođene ulazu u slikovnom formatu. Ipak, najviša srednja prosječna preciznost dobivena je u slučaju modificirane arhitekture ResNet-18 prilagođene radu sa sirovim podacima. Što se tiče inicijalizacije težina, rezultati pokazuju da korištenje skupa ImageNet u većini slučajeva pospješuje rezultat, dok korištenje skupa podataka LabSAR nema previše utjecaja osim kad se koristi u kombinaciji sa skupom ImageNet za modificiranu arhitekturu ResNet-18. Prema prikazanim rezultatima, u nastavku istraživanja fokusirali smo se na standardnu arhitekturu ResNet-18 za pristup rekonstruiranim slikama i modificiranu za pristup sirovim radarskim podacima.

**Tablica 5.3:** Usporedba rezultata svih razvijenih modela u kombinacijama s različitim inicijalizacijama težina. Prikazana je mjera srednje prosječne preciznosti (mAP).

Vrsta	Model	Inicijalizacija težina			
		Nasumično	ImageNet	LabSAR	ImageNet i LabSAR
RAW	MNv3	96.23	96.59	96.31	96.39
	RN18	97.12	97.42	97.27	97.38
	RAW-RN18	99.73	99.71	99.67	<b>99.95</b>
IMG	MNv3	98.29	98.71	98.34	98.66
	RN18	99.35	<b>99.75</b>	99.04	99.64

Kako bi se analizirao utjecaj veličine skupa za treniranje na rezultate klasifikacije, modeli su trenirani na podskupovima veličine 25, 50 i 75% originalnog skupa za treniranje. Još su jednom testirani utjecaji prije navedenih načina inicijalizacije težina modela. Tablica 5.4 prikazuje rezultate modela oba pristupa temeljena na dvije arhitekture koje su se pokazale najboljima, različitim načinima postavljanja početnih vrijednosti i veličinama skupa za treniranje.

Model temeljen na modificiranoj ResNet-18 arhitekturi i treniran sirovim radarskim poda-

**Tablica 5.4:** Usporedba rezultata najboljih modela za RAW i IMG pristup u kombinacijama s različitim načinima inicijalizacije težina u ovisnosti o veličini skupa za treniranje. Prikazana je mjera srednje prosječne preciznosti (mAP).

Model	Veličina skupa za treniranje			
	0.25	0.5	0.75	Full
IMG, RN18, Nasumično	96.56	98.72	99.33	99.35
IMG, RN18, ImageNet	96.25	98.63	99.33	<b>99.75</b>
IMG, RN18, LabSAR	98.07	99.65	99.03	99.04
IMG, RN18, ImageNet LabSAR	96.56	98.94	99.31	99.64
RAW, ModRN18, Nasumično	98.25	99.13	99.66	99.73
RAW, ModRN18, ImageNet	98.66	99.51	99.60	99.71
RAW, ModRN18, LabSAR	98.74	98.89	99.23	99.67
RAW, ModRN18, ImageNet i LabSAR	98.45	99.63	99.82	<b>99.95</b>

cima pokazao se otpornijim na smanjenje skupa za treniranje. Važno je napomenuti kako su prikazani rezultati dobiveni korištenjem najbolje epohe na skupu za validaciju što je standardni način testiranja robusnosti. Nakon što su na tom skupu za svaki pristup odabrani najbolji modeli, testirani su skupom za testiranje. U tablici 5.5 prikazane su prosječne preciznosti po klasama modela ResNet-18 arhitekture pretreniranog skupom ImageNet temeljenog na rekonstruiranim slikama (IMG model) i modela modificirane ResNet-18 arhitekture pretreniranog kombinacijom ImageNet i LabSAR skupa temeljenog na sirovim radarskim podacima (RAW model).

**Tablica 5.5:** Rezultati prosječne preciznosti (AP) po klasama IMG i RAW modela s kombinacijama u kojima su dali najbolje rezultate.

	Metal	Plastika	Staklo	mAP
IMG, ImageNet	98.07	97.42	98.07	97.85
RAW, ImageNet i LabSAR	98.71	98.12	97.85	98.23

Prosječna preciznost viša je u RAW modelu za klase koje predstavljaju metalni i plastični objekt, dok je IMG model bio nešto precizniji u klasifikaciji staklenog objekta. U prosjeku RAW model daje bolje rezultate od IMG modela.

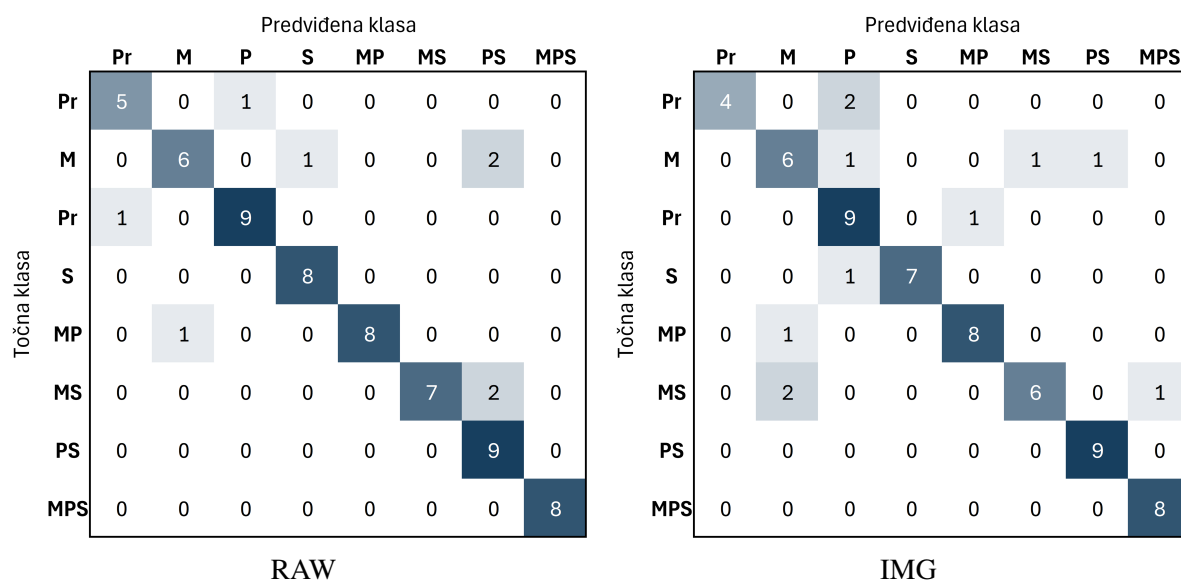
Kako bi se dodatno potvrdila ta tvrdnja, modeli su testirani u višeklasnoj klasifikaciji u kojoj svaka kombinacija objekata koji se mogu pojaviti u sceni čini jednu klasu. Tri testna objekta daju ukupno osam varijanti scena: prazna scena (Pr), tri scene s po jednim objektom (M, P,

S), tri s po dva objekta (MP, MS, PS) i jedna scena sa sva tri objekta prisutna (MPS). Osim najboljih modela za oba pristupa (RAW i IMG), analizirana je i klasifikacija modela čiji je izlazni vektor dobiven prosjekom vrijednosti izlaznih vektora RAW i IMG modela i nazvan je zajednički model (eng. *Ensemble Model*). Rezultati klasifikacije primjera iz skupa za testiranje RAW, IMG i zajedničkog modela prikazani su u tablici 5.6.

**Tablica 5.6:** Rezultati F1 mjere za svaku klasu za tri modela: IMG, RAW i zajednički (ENS) na test skupu. Klase su označene početnim slovima: Pr - prazna scena, M - metal, P - plastika, S - staklo, a kombinacije objekata kombinacijama slova. Na kraju je prikazana makro F1 mjera (mF1).

	Pr	M	P	S	MP	MS	PS	MPS	mF1
IMG	80.00	66.67	78.26	93.33	88.89	75.00	<b>94.74</b>	94.12	83.88
RAW	<b>83.33</b>	<b>75.00</b>	<b>90.00</b>	<b>94.12</b>	<b>94.12</b>	<b>87.50</b>	81.82	<b>100.00</b>	<b>88.24</b>
ENS	81.12	71.12	82.55	93.33	91.25	79.73	86.74	97.24	85.39

Može se primijetiti da su sve klase osim klase 'PS' što predstavlja scenu s plastičnim i staklenim objektom bolje klasificirane RAW modelom. U matricama zabune ta dva modela (slika 5.4) prikazan je razlog tako lošeg F1 rezultata za tu klasu. Naime, RAW model je pola pogrešno klasificiranih primjera klasificirao upravo tom klasom što, prema izrazu 3.18, snižava F1 rezultat.



**Slika 5.4:** Matrice zabune RAW i IMG modela. Klase su označene početnim slovima: Pr - prazna scena, M - metal, P - plastika, S - staklo, a kombinacije objekata kombinacijama slova.

Na matricama zabune može se vidjeti da od ukupno 67 testnih primjera, RAW model ih je pogrešno klasificirao 8 (88.06% točnost predikcije), a IMG 11 (83.58 %). Zanimljivo je i da je preklapanje tih primjera stopostotno, odnosno da nema nijednog testnog primjera koji je

IMG model ispravno klasificirao, a RAW model pogrešno. Takav rezultat implicira da rekonstrukcijski algoritam, što je i očekivano, ne može dodati novu informaciju koja bi bila značajna modelu za razlikovanje objekata. Kao što je opisano u poglavlju 2.2.1, algoritam za rekonstrukciju radarske slike kroz niz koraka pretvara sirove radarske podatke u sliku koju je čovjeku lakše interpretirati. U idealnom slučaju u tim se koracima ne narušava integritet podataka, ali u stvarnosti se zbog aproksimacija i interpolacije na kraju, gubi dio informacije. Budući da ne postoji primjer koji je ispravno klasificiran u IMG modelu, a pogrešno u RAW, za očekivati je da zajednički model neće imati bolje rezultate. Glavna ideja takvog modela je da se klasifikacija rubnih primjera specifičnih jednom modelu olakša kombiniranjem izlaza tog modela s drugim. Kako IMG model nijedan primjer nije bolje klasificirao, zajednički model nije mogao imati bolji rezultat od RAW modela.

Što se tiče materijala objekata, matrica zabune pokazuje da su oba modela imala problema s klasifikacijom scena u kojima se pojavljuje metalni objekt. Razlog za to je veća refleksija takvog objekta od staklenog i plastičnog zbog koje se u sirovim podacima povećava omjer signala dobivenih refleksijom od tog objekta i signala dobivenih refleksijom od druga dva. U pristupu s rekonstruiranim slikama taj se problem dodatno naglašava kroz prikaz objekata različitim intenzitetima. Naime, ako se u vizualizaciji rekonstruirane slike toplinskom mapom ograniči skala intenziteta, to može rezultirati ili premalo naglašenim objektom (u slučaju metalnog objekta) ili neprikazivanjem objekta (u slučaju plastičnog objekta). S druge strane, ne postavljanje ograničenja znači da pikseli iste boje ne predstavljaju isti intenzitet što opet može imati posljedice na klasifikaciju. Iako rezultati potvrđuju da mreža bolje uči na sirovim radarskim podacima (RAW model) nego na obrađenima (IMG model), čovjeku je lakše interpretirati rekonstruirane slike pa je samim tim i identifikacija uzroka zašto je neki primjer ispravno ili pogrešno klasificiran jednostavnija na njima. S druge strane, za pristup klasifikaciji sirovim podacima nije potrebno provoditi rekonstrukcijski algoritam što može biti značajno kod implementacije u ugradbenim računalima s ograničenim resursima. Skup podataka LabSAR na primjeru RAW modela pokazao je da se takvi i slični podaci mogu primijeniti u inicijalizaciji težina modela.

## **5.2 Modeli trenirani podacima snimanim različitim radarskim parametrima**

Modeli korišteni u usporedbi temelje se na modificiranom RAW modelu iz poglavlja 5.1.2, a skup podataka je opisan u 4.3. Taj je skup omogućio analizu utjecaja radarskih parametara na rezultat klasifikacije. Prilikom mjerenja mijenjani su parametri duljina koraka i širina frekventijskog pojasa FMCW signala kako bi se na rezultatnoj matrici utjecalo u obje osi – onoj koja predstavlja prostor kojim se radar kreće i onoj koja predstavlja udaljenost. Povećanjem duljine koraka isti se prostor snima s manje pozicija što samo po sebi smanjuje količinu informacije

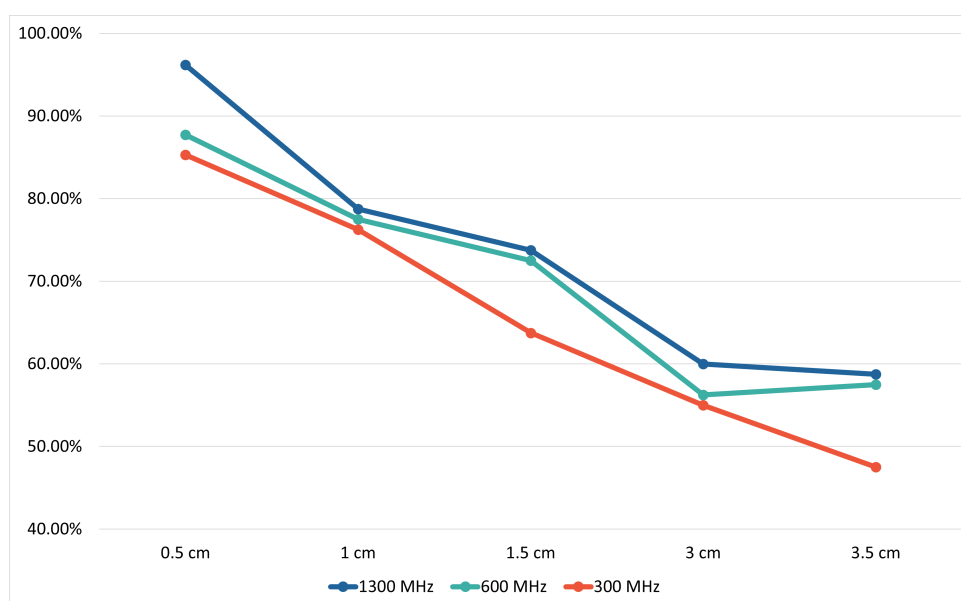


o sceni, a promjenom širine pojasa izravno se utječe na rezoluciju u smjeru udaljenosti. Veća širina pojasa unutar kojeg se frekvencija chirp signala linearno povećava može uhvatiti veće detalje o promatranom objektu kao što su oblik i struktura površine.

Provedene su dvije analize. U prvoj su modeli trenirani istim brojem primjera što omogućava smislenu usporedbu rezultata. U drugoj su analizi modeli trenirani svim dostupnim primjerima što znači da su modeli s većom duljinom koraka trenirani s više primjera nego oni s manjom duljinom koraka. Takav pristup ograničava izravnu usporedbu rezultata, ali može prikazati kako bi se modeli ponašali kad bi se povećao skup podataka.

### 5.2.1 Analiza na istom broju primjera

Prva je analiza temeljena na 81 primjeru budući da je toliko originalnih primjera snimljeno duljinom koraka 0.5 cm, a od kojih su kasnije generirane matrice za veće duljine koraka. Za svaku od pet duljina koraka (0.5, 1, 1.5, 3, 3.5 cm) kreirana su tri skupa podataka (po jedan za svaku širinu pojasa) koji su korišteni u treniranju modela. Distribucija klasa svakog od korištenih skupova je jednaka distribuciji u originalnom skupu. Ukupno je provedeno pet iteracija treniranja i testiranja različitim primjerima svih modela. Na slici 5.5 prikazane su prosječne točnosti modela kroz tih pet iteracija za svaku kombinaciju duljine koraka i širine pojasa.



**Slika 5.5:** Prosječna točnost klasifikacije modela treniranog podacima dobivenim različitim kombinacijama duljine koraka i širine pojasa. U treniranju, validaciji i testiranju korišten je jednak broj primjera (81) [95].

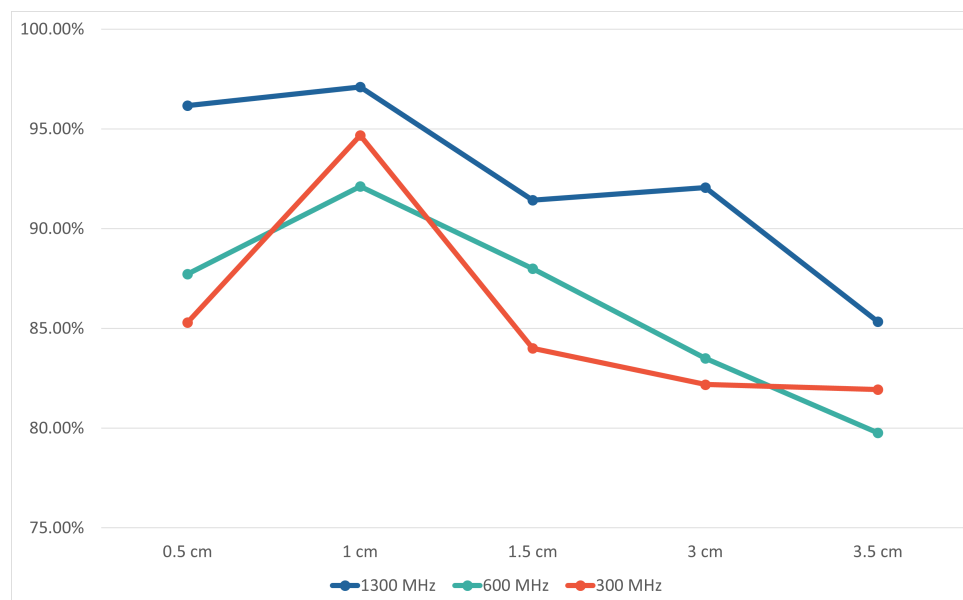
Rezultati pokazuju smanjenje točnosti kad se koristi uža širina pojasa što je i očekivano budući da se time smanjuje i količina informacija i samim time točnost modela. Na slici 5.5 može se vidjeti da u svim kombinacijama s različitim duljinama koraka veća širina pojasa ima veću točnost klasifikacije. Konkretno u slučaju duljine koraka od 0.5 cm, točnost s 96.17% za

1300 MHz širine frekvencijskog pojasa pada na 87.73% za 600 MHz, odnosno 85.29% za 300 MHz.

Slično se ponašanje - pad točnosti klasifikacije - može primijetiti i u slučaju povećanja duljine koraka. Kad promatramo rezultate zadržavajući istu širinu pojasa (na primjeru 1300 MHz), točnost je u svakoj iteraciji modela treniranog na podacima dobivenih duljinom koraka od 0.5 cm bila viša od 93.75%, a u dvije iteracije je dosegla stopostotnu točnost klasifikacije. S druge strane, za istu širinu pojasa, a dvostruko veću duljinu koraka (1 cm), najviša točnost u svih pet iteracija bila je 81.25%, dok je u najzahtjevnijem slučaju (duljina koraka 3.5 cm) točnost pala na vrijednosti između 56.25% i 32.5%.

## 5.2.2 Analiza na drugačijem broju primjera

U drugom su setu iteracija testirani modeli koji su za različite duljine koraka imali različit broj primjera u skupu podataka. Budući da su podaci koji predstavljaju mjerenja dobivena većim duljinama koraka umjetno generirani iz originalnih (kod kojih je korištena duljina koraka 0.5 cm), takvih je podataka više nego u originalnom skupu. Tako su primjerice od 81 matrice dobivene originalnim mjerenjem kreirane 162 matrice koje predstavljaju mjerenja dobivena duljinom koraka 1 cm. Cilj je ove analize pokazati potencijal klasifikacije objekata radarskim podacima kad na raspolaganju imamo veći skup. Prosječne točnosti prikazane su na slici 5.6.



**Slika 5.6:** Prosječna točnost klasifikacije modela treniranog podacima dobivenim različitim kombinacijama duljine koraka i širine pojasa. U treniranju, validaciji i testiranju korišten je različiti broj primjera. Modeli s većom duljinom koraka imaju više primjera [95].

Rezultati pokazuju povećanje točnosti u svim slučajevima (osim kod originalnih mjerenja) u odnosu na prvu analizu u kojoj je ograničen broj primjera. To nije iznenađenje budući da u svim takvim skupovima modeli imaju više primjera za učenje, validaciju i testiranje. Točnosti iznad

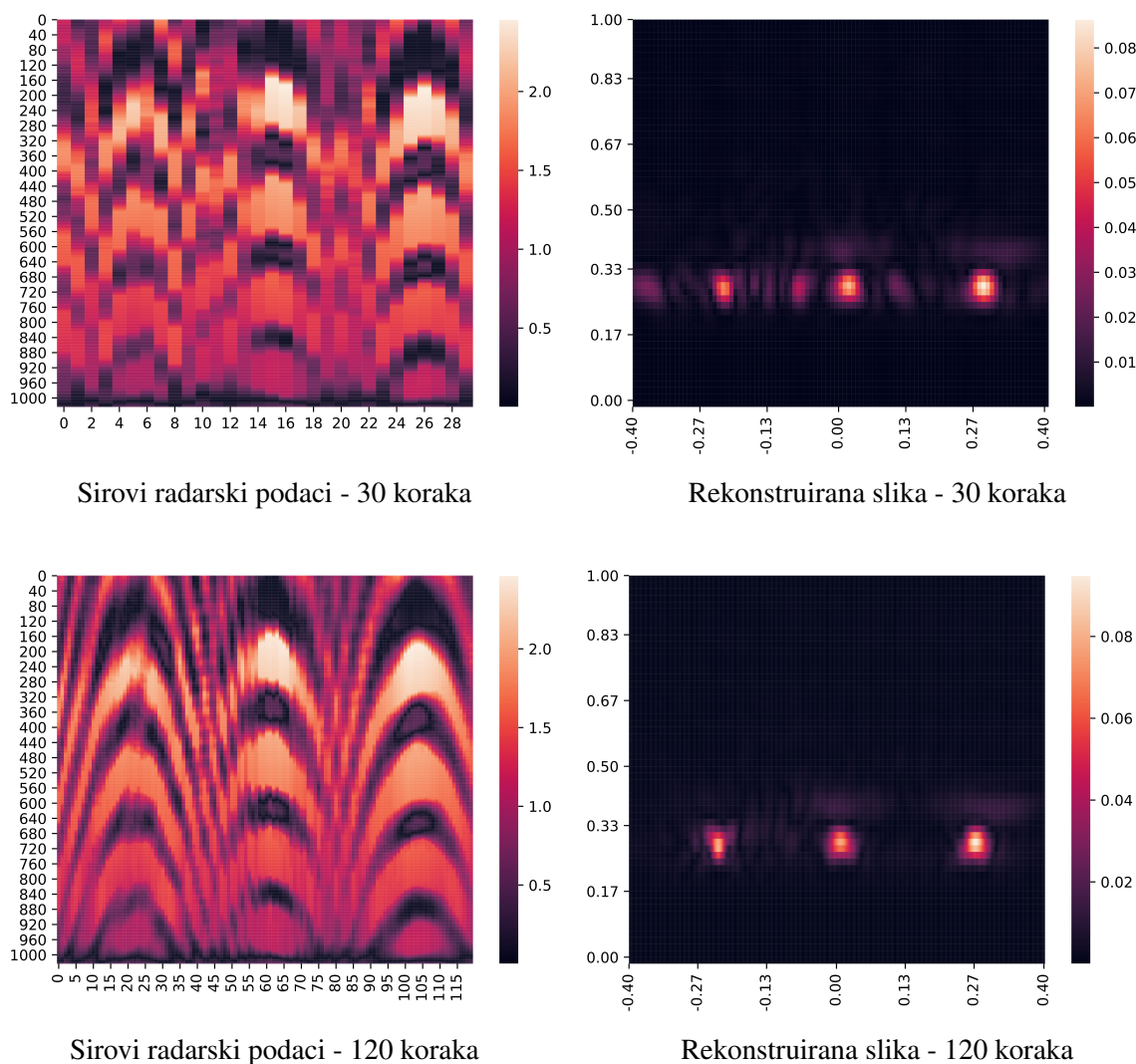
90% zabilježene su u svim slučajevima širine pojasa 1300 MHz osim u kombinaciji te širine pojasa s najduljim korakom (3.5 cm) iako je čak i s tako izazovnom duljinom koraka (sedam puta većoj od originalne), točnost varirala između 81.01% i 87.8%. Pogoršanje se može primijetiti u slučajevima uže širine pojasa u kojima je točnost pala na raspon 80-90%. Zanimljivo je da su najviše razine točnosti zabilježene u slučaju duljine koraka od 1 cm, a ne 0.5 cm. Na tim se rezultatima najbolje vidi utjecaj većeg skupa podataka. Iako su prosječne točnosti za te dvije duljine koraka u kombinaciji s najvećom širinom pojasa (1300 MHz) usporedive, za uže širine pojasa razlika postaje sve veća. Tako je model treniran skupom kombinacije 0.5 cm i 300 MHz imao prosječnu točnost od 86.54%, a 1 cm i 300 MHz 94.68%. Rezultati impliciraju da bi se uz veći skup podataka mogla povećati udaljenost između dva uzastopna mjerenja na GBSAR traci (duljina koraka) čime bi se smanjila potrošnja energije, a i dalje zadržati zadovoljavajuću točnost klasifikacije. Upravo se na tome temelji ideja sljedećeg modela.

### 5.3 Klasifikacija objekata s optimizacijom parametara

Razvijeni GBSAR pomiče FMCW radar po traci u stani-kreni (eng. *stop-and-go*) načinu rada. U prošlom je pristupu pokazano da više koraka, odnosno pozicija s kojih se promatra neka scena daje više informacije i povećava točnost klasifikacije. S druge strane, zaustavljanje senzora i snimanje s više pozicija dodatno troši energiju. Ideja ovog pristupa je pronaći balans između potrošnje energije i točnosti klasifikacije.

Pristup je testiran na skupu za klasifikaciju s 9 objekata opisanom u poglavlju 4.4. U skupu su snimane scene s tri objekta koje su kasnije podijeljene na tri segmenta s jednim objektom. Primjeri iz skupa su dodatno obrađeni na način da se matricama dimenzija 1024x40 koje predstavljaju jedan originalni segment uklanjanjem tri od četiri uzastopna mjerenja generira drugi skup koji predstavlja isti segment sniman iz manje pozicija, odnosno s većom duljinom koraka. Generirani skup s matricama dimenzija 1024x10 nazvan je inicijalni, a originalni (s matricama 1024x40) završni. Primjer sirovih podataka i rekonstruirane matrice cijele scene prikazan je na slici 5.7. U gornjem redu su podaci snimani u 30 koraka, odnosno 10 po segmentu, a u donjem sa 120 koraka što daje 40 po segmentu. Na slici se vidi kako je rezolucija rekonstruirane slike snimane sa 120 pozicija viša što je i očekivano budući da ima četiri puta više podataka.

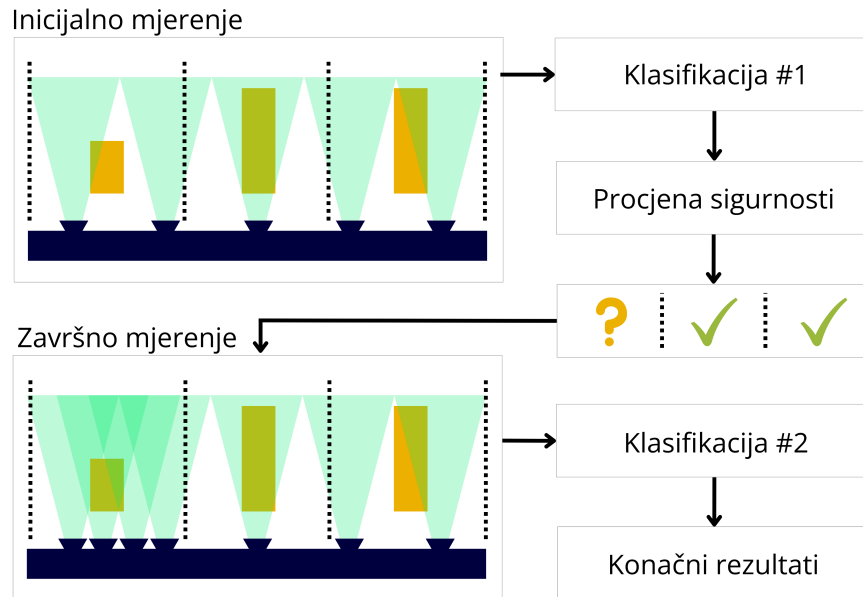
Glavna ideja ovog pristupa je iskoristiti komunikaciju između radara i modela. Konkretnije, cilj je snimiti cijelu scenu s tri objekta s manje pozicija (u našem slučaju u 30 koraka), podijeliti snimljenu matricu na tri jednaka dijela (segmenti inicijalnog skupa dimenzija 1024x10) od kojih se u svakom nalazi jedan objekt i klasificirati objekte. Nakon klasifikacije za svaki se primjer određuje sigurnost modela da je predviđena klasa upravo ta koja treba biti, a ako je sigurnost modela ispod određene razine radar ponavlja snimanje tog segmenta s dodatnih pozicija – u našem slučaju s dodatnih 30 pozicija (segmenti završnog skupa dimenzija 1024x40).



**Slika 5.7:** Primjeri sirovih radarskih podataka i rekonstruiranih slika snimljeni u 30 i 120 koraka. Promatrana scena ima tri metalna objekta različitih oblika.

U implementaciji ideje za sigurnost modela korištena je funkcija softmax primijenjena na izlaz modela. Softmax za svaku klasu daje vjerojatnost da određeni primjer predstavlja upravo tu klasu. Blokovski prikaz algoritma koji koristi 'povratnu vezu' između modela i radara prikazan je na slici 5.8.

Razvijena su dva modela: inicijalni koji je treniran na inicijalnom skupu matricama dimenzija  $1024 \times 10$ , a koje predstavljaju sirove radarske podatke snimane iz 10 pozicija i završni treniran na završnom skupu matricama dimenzija  $1024 \times 40$ , a koje predstavljaju sirove radarske podatke snimane iz 40 pozicija, odnosno 10 iz inicijalnog i 30 dodatnih. Važno je napomenuti da je u oba slučaja prostor istih dimenzija pokriven radarom. Oba su modela trenirana i testirana kroz pet iteracija tijekom kojih su primjeri skupova za treniranje, validaciju i testiranje nasumično raspodijeljeni. U svakoj su iteraciji isti primjeri bili u ta tri skupa za oba modela tako da se točnosti mogu izravno usporediti. Rezultati se mogu promatrati kroz prije navedenu ideju



**Slika 5.8:** Prikaz algoritma s povratnom vezom koji omogućava komunikaciju između modela dubokog učenja i radara s ciljem prilagodbe parametara radara za detaljnije snimanje određenog dijela scene [96].

određivanja sigurnosti i snimanja s dodatnih pozicija. U njoj funkcija softmax nakon inicijalnog modela određuje vjerojatnosti za svaku klasu. Ako je vjerojatnost za neku klasu iznad određene granice, za taj se segment određuje ta klasa. S druge strane, ako vjerojatnost ni za jednu klasu nije iznad postavljene granice, radar ponavlja mjerenje tog segmenta s dodatnih pozicija. Kako bi se testirao predloženi koncept, prema početnim rezultatima postavljena je granica na 80%. To znači da se svim segmentima kojima nijedna predviđena klasa nema vjerojatnost dobivenu funkcijom softmax na izlazu inicijalnog modela iznad 80% uzima rezultat završnog modela treniranog na skupu koji predstavlja mjerenja dobivena s 40 pozicija. U tablici 5.7 prikazane su točnosti svake iteracije prvo za inicijalni model, zatim broj segmenata kojima je ponovljeno snimanje s dodatnih pozicija i na kraju konačni rezultat dobiven modelom treniranom na završnom skupu. Rezultati pokazuju da je u svakoj iteraciji povećana konačna točnost u odnosu na inicijalnu. Broj segmenata kojima je ponovljeno mjerenje nakon inicijalnog modela je varirao između 9 i 20 primjera od ukupno 54 testna primjera po iteraciji. Konkretnije, od ukupno 270 testnih primjera (5 iteracija po 54 primjera) ponovljeno je njih 73, odnosno 27%. To znači da je 73% testnih primjera klasificirano sa samo 10 pozicija dok je za ostale trebala informacija s dodatnih 30 pozicija. Konačna točnost je u svim iteracijama bila iznad 90%. Najveći porast završne klasifikacije u odnosu na inicijalnu dogodio se u četvrtoj iteraciji kad je s 81.48% točnosti nakon 20 ponovljenih primjera točnost skočila na 92.59%. U petoj iteraciji umjesto snimanja svih 54 primjera s 40 koraka, njih 45 snimano je s 10, a preostalih 9 sa svih 40 što je ukupno 810 FMCW mjerenja u odnosu na 2160 (54 primjera po 40 koraka). Odnosno, u toj je iteraciji, na našem primjeru, za klasifikaciju objekata točnošću 90.74% bilo dovoljno 62.5%

manje mjerenja u odnosu na pristup sa snimanjem iz svih pozicija.

**Tablica 5.7:** Točnosti RAW modela po iteracijama na inicijalnom i završnom mjerenju.

Iteracija	Točnost nakon inicijalnog mjerenja	Broj primjera ispod praga	Konačna točnost
1	81.48%	18	90.74
2	88.88%	11	92.59
3	83.33%	15	92.59
4	81.48%	20	92.59
5	87.04%	9	90.74

### 5.3.1 Određivanje sigurnosti modela

Budući da izlaz funkcije softmax nije preporučeno koristiti kao metriku za sigurnost modela jer ono predstavlja vjerojatnost za neku klasu da je ulazni primjer upravo ta klasa, postoji mogućnost da je ono što u ovoj analizi nazivamo sigurnost modela pogrešno određeno. Kao potencijalni problem u tom se slučaju ističu testni primjeri izvan skupa koji mogu dobiti visoku vjerojatnost jer ne sliče nijednom primjeru iz seta za treniranje što kod nas u fazi testiranja koncepta neće biti slučaj. Ipak, kako bismo testirali koliko dobro vjerojatnosti dobivene funkcijom softmax predstavljaju očekivanu sigurnost, na predikcijama svih iteracija analizirali smo koliko su često ispravno klasificirani primjeri čije su predviđene vjerojatnosti upale u sljedeće grupe: 0-50%, 50%-60%, 60%-70%, 70%-80%, 80%-90% i 90%-100%. Cilj je da primjeri koje model klasificira s vjerojatnošću između 80% i 90% u 80 do 90 % slučajeva budu točno klasificirani. Odnosno, da se vjerojatnost poklapa s učestalosti točne predikcije. Omjeri točno klasificiranih i ukupno primjera za svaku od grupa dani su u tablici 5.8. Može se primijetiti da je samo u slučaju primjera s predviđenom vjerojatnošću između 70 i 80% u više primjera točno klasificirano nego što je model to predvidio. Za sve ostale grupe vjerojatnost je ispravno određivala točnost klasifikacije što implicira da je u ovom jednostavnom slučaju funkcija softmax dobro predstavljala sigurnost modela.

Rezultati prikazani u tablici 5.7 pokazali su potencijal pristupa u kojem komunikacija između klasifikacijskog modela i radara omogućava prilagodbu radarskih parametara s ciljem uštede energije i bolje klasifikacije. Osim duljine koraka kroz takvu se komunikaciju mogu postavljati i ostali parametri – npr. polarizacija. Budući da je cijeli sustav koji objedinjuje klasifikacijski model i radar razvijen na istom mikroročunalu, komunikacija između njih ostvarena je programski. Upravo je to bio razlog implementacije cijelog sustava na jednom mikroročunalu.

Valja napomenuti i kako pristup u ovakvoj izvedbi GBSAR-a u kojoj se senzor kreće po

**Tablica 5.8:** Postotak ispravno klasificiranih primjera grupe

<b>Grupa</b>	<b>Ispravno klasificirano</b>
0-50	26.47 %
50-60	59.09 %
60-70	66.00 %
70-80	84.13 %
80-90	88.09 %
90-100	94.39 %

vodilici samo donekle ispunjava energetska učinkovitost jer senzor ipak prelazi cijelim putem kao i u slučaju s više koraka tako da se ušteda ostvaruje zbog manje zaustavljanja i FMCW mjerenja. S druge strane, u izvedbi GBSAR-a s nizom statičnih senzora raspoređenih po traci koje radar može, a ne mora koristiti prilikom snimanja, sustav može ostvariti značajnije uštede korištenjem manjeg broja senzora.

## 5.4 Klasifikacija na skupovima podataka s različitim polarizacijama

U prošlim je poglavljima analiziran utjecaj širine frekvencijskog pojasa i duljine koraka na klasifikaciju. U ovom se analiziraju rezultati modela trenirani na skupovima dobivenim radarskim snimanjem s različitim polarizacijama. U dva su uzastopna mjerenja sve scene snimane prvo horizontalnom, a onda vertikalnom polarizacijom. Korišteni skup podataka detaljnije je opisan u poglavlju 4.5. Osim originalnih skupova koji sadrže matrice snimljene horizontalnom (skup HH) i vertikalnom (skup VV) polarizacijom, kombiniranjem elemenata tih matrica kreirani su dodatni skupovi:

- SUB – skup dobiven oduzimanjem matrice snimane vertikalnom polarizacijom od one snimane horizontalnom polarizacijom
- AVG – skup dobiven usrednjivanjem matrica snimanih vertikalnom i horizontalnom polarizacijom
- MIX\_ROWS – skup dobiven naizmjeničnim miješanjem redova matrica snimanih horizontalnom i vertikalnom polarizacijom
- MIX\_COL – skup dobiven naizmjeničnim miješanjem stupaca matrica snimanih horizontalnom i vertikalnom polarizacijom
- JOIN – skup dobiven dodavanjem matrice snimane vertikalnom polarizacijom na kraj

matrice snimane horizontalnom polarizacijom.

Različite kombinacije sadrže iste podatke - elemente matrica dobivenih horizontalnom i vertikalnom polarizacijom, prikazane na različite načine. Ideja je testirati koji od navedenih načina najviše odgovara klasifikacijskom modelu, odnosno treniranje nad kojim će skupom podataka ostvariti najvišu točnost. Dimenzije matrica navedenih skupova dane su u tablici 5.9, a na slici 5.9 prikazane su toplinske mape primjera iz svakog skupa. Ti su primjeri dobiveni kombinacijama matrica snimanih horizontalnom i vertikalnom polarizacijom sa slike 4.4.

**Tablica 5.9:** Dimenzije matrica generiranih skupova. 'HH' označava matricu snimljenu horizontalnom polarizacijom, a 'VV' vertikalnom.

HH	VV	SUB	AVG	MIX ROWS	MIX COL	JOIN
1024x20	1024x20	1024x20	1024x20	2048x20	1024x40	1024x40

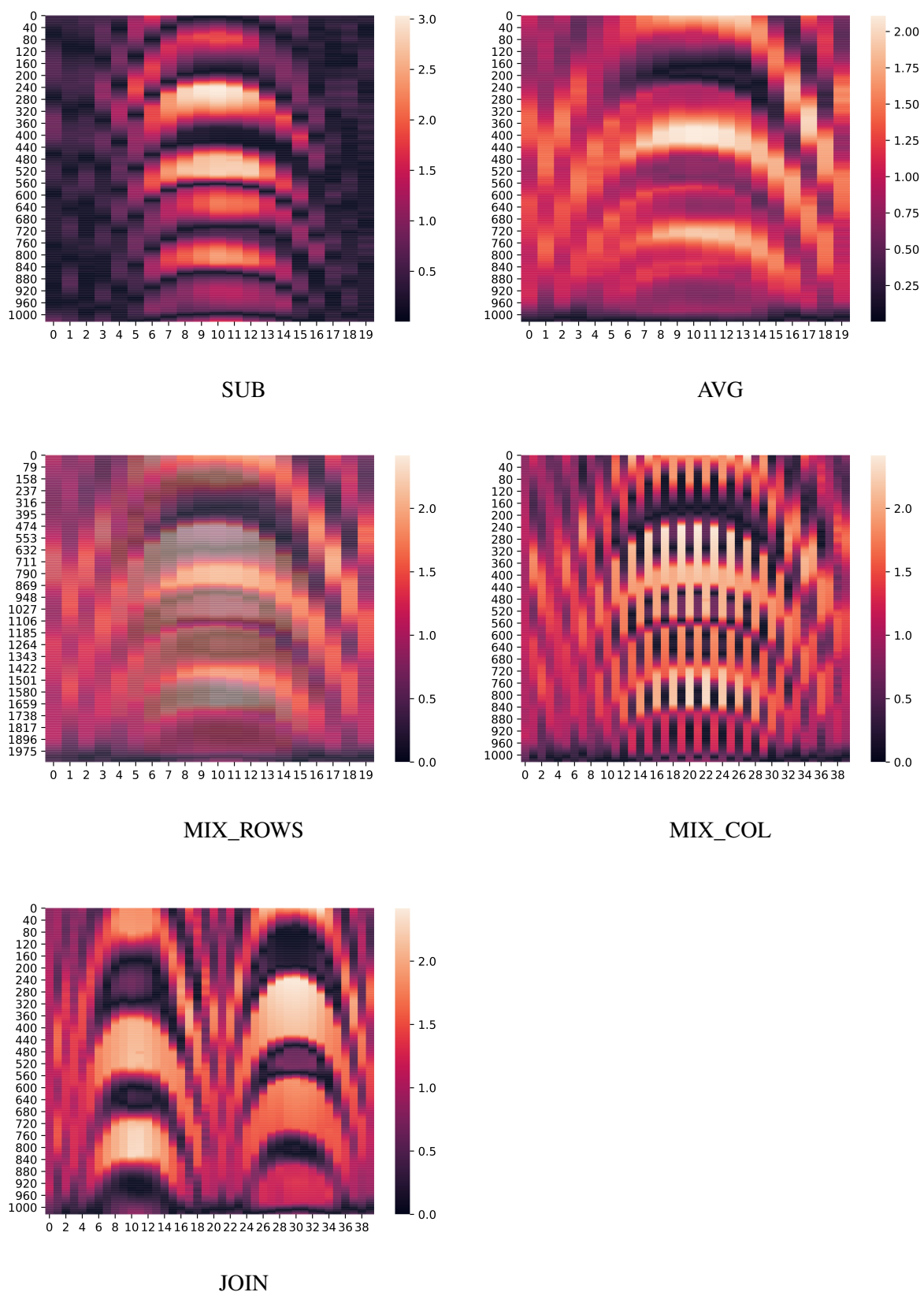
Svakim je skupom treniran jedan model što znači da je ukupno razvijeno sedam modela koji su testirani kroz 10 iteracija. Tijekom svake od iteracija modeli su trenirani, validirani i testirani na istim primjerima, odnosno na njihovim pandanima druge polarizacije. Dakle, ako je model horizontalne polarizacije testiran na primjeru 'AS\_3\_hh\_48' koji označava metalni objekt pravokutnog oblika sniman horizontalnom polarizacijom na trećoj poziciji u sceni 48. mjerenja, među testnim primjerima te iteracije za vertikalnu polarizaciju naći će se 'AS\_3\_vv\_48', a za ostale primjer 'AS\_3\_48' koji je dobiven kombinacijama 'AS\_3\_hh\_48' i 'AS\_3\_vv\_48'.

### 5.4.1 Rezultati modela trenirani na kreiranim skupovima

U tablici 5.10 (i na slici 5.10) prikazane su najniže, najviše i prosječne točnosti 10 iteracija za svaki model. Može se primijetiti da su modeli trenirani podacima dobivenim horizontalnom polarizacijom u prosjeku bili točniji od onih treniranih podacima vertikalne polarizacije (89.43% u odnosu na 75.71%). U istoj su iteraciji oba modela stopostotnom točnošću odredili klase svih testnih primjera, ali je u 8 od 9 preostalih iteracija model treniran na HH skupu bolje klasificirao. Razlog može biti odabir testnih objekata i njihova pozicija, odnosno orijentacija u sceni, ali i hardverska ograničenja korištenog FMCW modula čiji dijagram zračenja, prema tehničkim karakteristikama, pokazuje drugačije ponašanje u horizontalnoj i vertikalnoj osi.

Što se tiče kombinacija, najslabiji je rezultat postigao model treniran SUB skupom u kojem su matrice oduzete jedna od druge. To je i očekivano budući da se na taj način gubi dio informacije u procesu oduzimanja. Na sličan se način može gledati i AVG skup u kojem se dio informacije gubi usrednjavanjem matrica. Druga skupina kombinacija (MIX\_COL, MIX\_ROWS i JOIN) sastoji se od svih elemenata obiju originalnih matrica (horizontalne i vertikalne polarizacije) pa je samim time očekivano da će dati bolje rezultate od prve dvije kombinacije. Dakle,

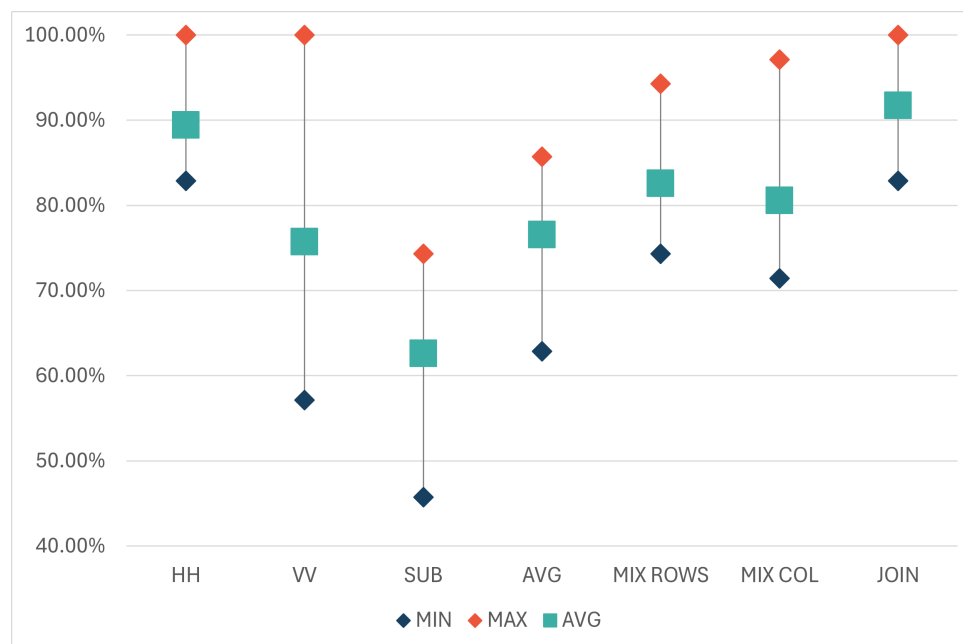




**Slika 5.9:** Primjeri sirovih radarskih podataka kreiranih skupova: SUB, AVG, MIX\_ROWS, MIX\_COL i JOIN. Generirani primjeri nastali su kombinacijom primjera prikazanih na slici 4.4 [94]

**Tablica 5.10:** Najmanja, najveća i srednja postignuta točnost iskazana u postotku [%] kroz 10 iteracija provedenih na modelima treniranim skupovima HH, VV, SUB, AVG, MIX\_ROWS, MIX\_COL i JOIN.

	HH	VV	SUB	AVG	MIX ROWS	MIX COL	JOIN
Min.	82.86	57.14	45.71	62.86	74.29	71.43	82.86
Max.	100	100	74.29	85.71	94.29	97.14	100
Prosjek	89.43	75.71	62.6	76.57	82.57	80.58	91.72



**Slika 5.10:** Najmanja, najveća i prosječna točnost modela HH, VV, SUB, AVG, MIX\_ROWS, MIX\_COL, i JOIN kroz prvi set od 10 iteracija testiranja [94].

u sva tri skupa korišteni su isti, ali drugačije raspoređeni elementi matrice. Zanimljivo je da modeli sa skupovima MIX\_COL i MIX\_ROWS daju slične rezultate (prosjek prvog je 80.58%, a drugog 82.57%), dok skup JOIN u kojem su originalne matrice postavljene jedna do druge daje najviše rezultate u usporedbi sa svim treniranim modelima (u prosjeku 91.72%).

Ovakvi se rezultati mogu promatrati iz perspektive izvan područja radarskih signala – kroz utjecaj redosljeda elemenata iste matrice na rezultat klasifikacije iste arhitekture konvolucijske neuronske mreže. Naime, u skupu JOIN redosljed elemenata je isti kao i u originalnim matricama pa je uz same vrijednosti zadržana i informacija o pozicijama i rasporedu pojedinih signala, dok to kod MIX\_COL i MIX\_ROWS nije slučaj. U njima su originalne matrice rascjepkane u stupce, odnosno redove. Potencijalni nedostatak skupa JOIN je to što su signali dobiveni na istoj poziciji horizontalnom i vertikalnom polarizacijom u JOIN matrici udaljeni za širinu jedne matrice (broj koraka GBSAR-a). Konkretno, u našem slučaju 1. stupac JOIN matrice sniman je s iste pozicije kao i 21. stupac, ali drugačijim polarizacijama. S druge strane,

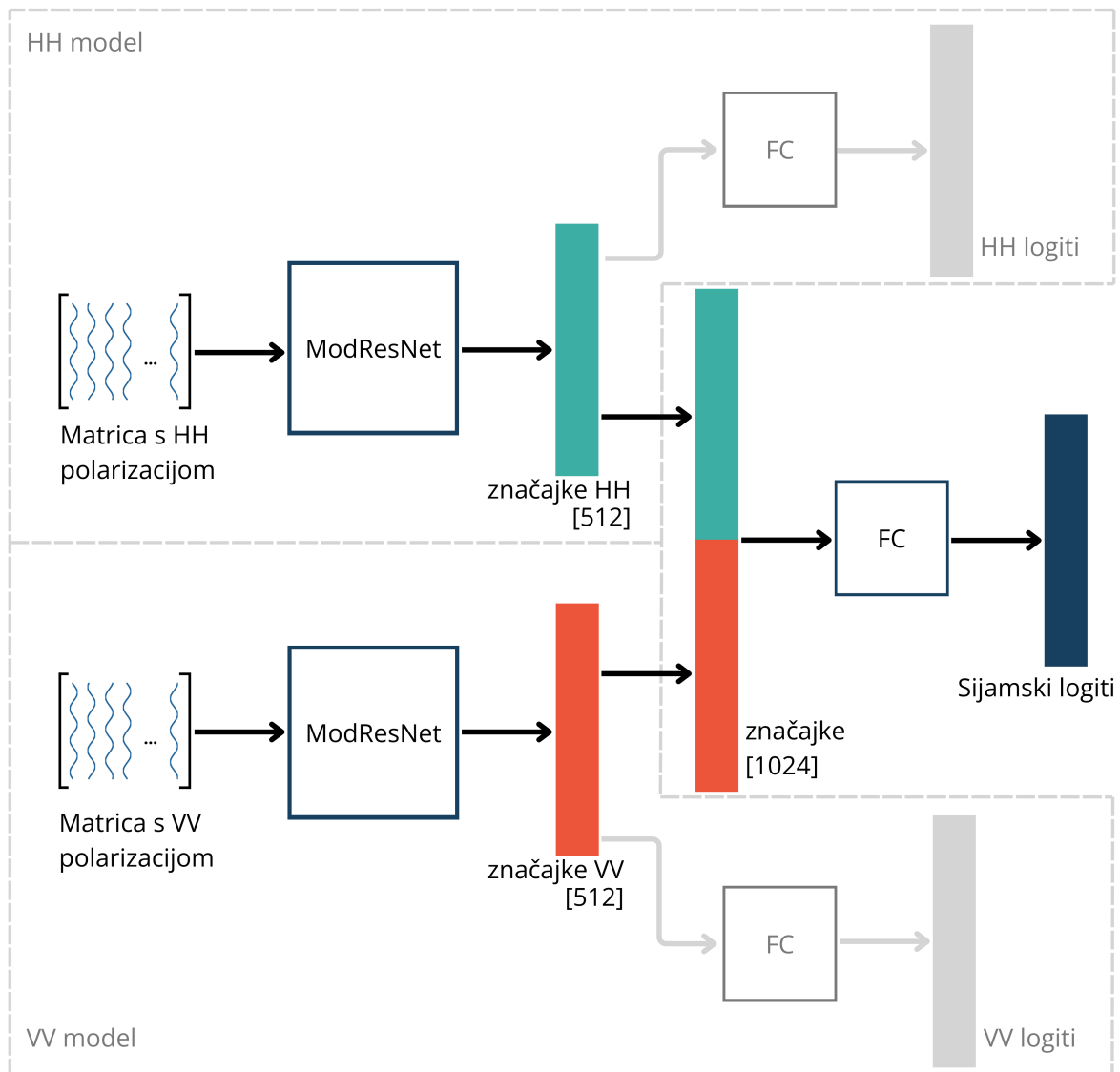
u MIX\_COL skupu dva susjedna stupca predstavljaju signale snimane s iste pozicije, ali je redosljed unutar dviju originalnih matrica rascjepkan što – pokazuju rezultati – narušava točnost klasifikacije. U MIX\_ROWS su također rascjepkane obje matrice, ali po redovima što opet smanjuje točnost klasifikacije. Skup JOIN koji daje najbolje rezultate, dakle, ima prednost očuvanja redosljeda unutar matrica, ali i problem gledanja dviju spojenih matrica kao jednu u kojoj udaljenosti između stupaca ne predstavljaju jednaku prostornu udaljenost. Kako bi se doskočilo tom nedostatku, analizirano je još nekoliko pristupa u drugom setu 10 iteracija.

### 5.4.2 Zajednički i Sijamski modeli

U tom su pristupu rezultati modela s matricama snimljenim horizontalnom (HH) i vertikalnom (VV) polarizacijom povezani u zajednički (eng. *Ensemble*) model u kojem prosjek softmax rezultata ta dva modela predstavlja konačni softmax rezultat za svaku klasu. Točnost zajedničkog modela, koji će u nastavku biti označen s 'ENS', izravno ovisi o rezultatima HH i VV modela. U slučaju kad oba modela na istim primjerima pogriješe, taj će primjer najčešće biti pogrešno klasificiran i u zajedničkom modelu. Ipak, primjećeno je nekoliko slučajeva u kojima se dogodilo suprotno. U jednom od njih model s horizontalnom polarizacijom klasificirao je objekt klase 'GR' kao 'AR', a model s vertikalnom kao 'GS'. Dakle, prvi je pogriješio u materijalu, ali je točno raspoznao oblik, dok je drugi prepoznao o kojem se materijalu radi, a pogriješio oblik. Ali, budući da su i jedan i drugi model kao drugu najvjerojatniju klasu predvidjeli upravo 'GR' i to s visokom vjerojatnošću, zajednički model je ispravno klasificirao taj primjer. Rezultati kroz svih 10 iteracija prikazani u tablici 5.11 potvrdili su višu točnost modela s horizontalnom polarizacijom u odnosu na onog s vertikalnom. U prosjeku je horizontalni model uspješno klasificirao 85.31% primjera testnog skupa, dok je vertikalni to učinio za 75.51% primjera. Zanimljivo je da je u prosjeku zajednički model nadmašio rezultate oba modela. Od 10 iteracija, zajednički model bio je točniji u 6, u 3 je izjednačio rezultat horizontalnog modela, a samo jednom je bio lošiji od njega. U toj je jednoj iteraciji vertikalni model snizio točnost zajedničkog. Zato je testirana još jedna varijanta zajedničkog modela. U njoj je, zbog superiornosti horizontalnog modela, kod uprosječivanja rezultata, vjerojatnosti klasa tog modela povećan utjecaj množenjem određenim faktorom većim od 1 kod računanja vjerojatnosti za zajednički model. Međutim, takav je pristup u samo jednoj iteraciji poboljšao rezultat zajedničkog modela, dok je u ostalim postigao istu točnost.

U drugom pristupu razvijen je model u kojem dvije grane klasifikacije uče paralelno. Svaka grana arhitekturom je slična arhitekturi modificiranog ResNet-18 modela i na ulazu prima matricu jedne polarizacije, ali na izlazu umjesto potpuno povezanog sloja daje vektor značajki te polarizacije koji se povezuje s vektorom značajki druge grane kreiranim na isti način, ali za drugu polarizaciju. Ta dva vektora, svaki veličine 512, zajedno čine vektor značajki cijelog modela (veličine 1024) koji ide prema potpuno povezanom sloju nakon čega se, slično kao u

originalnom modelu, računaju vjerojatnosti za svaku klasu i određuje predviđanje. Taj model je u nastavku označen sa 'Sijamski', a njegova arhitektura prikazana je na slici 5.11.



**Slika 5.11:** Arhitektura Sijamskog modela koji paralelno uči dvije grane za matrice snimljene različitim polarizacijama. Grane spajaju svoje vektore značajki prije potpuno povezanog sloja koji daje jedinstveni izlaz modela [94].

Na taj način model može zasebno trenirati dio mreže za svaku polarizaciju i na kraju ih povezati kako bi učenjem na obje matrice odredio klasu. Iako je takav pristup u prosjeku dao nešto bolje rezultate nego modeli trenirani na jednoj polarizaciji, nije dobiveno značajnije poboljšanje što je suprotno očekivanom zbog dvostruko više podataka i složenije mreže. Valja napomenuti kako je skup podataka korišten u svim ovim modelima relativno mali što, posebno kod složenijih mreža, može negativno utjecati na klasifikaciju zbog nedovoljnog skupa za treniranje kojima se postavljaju težine. Zato smo testirali isti pristup s postavljenim težinama za oba pod-

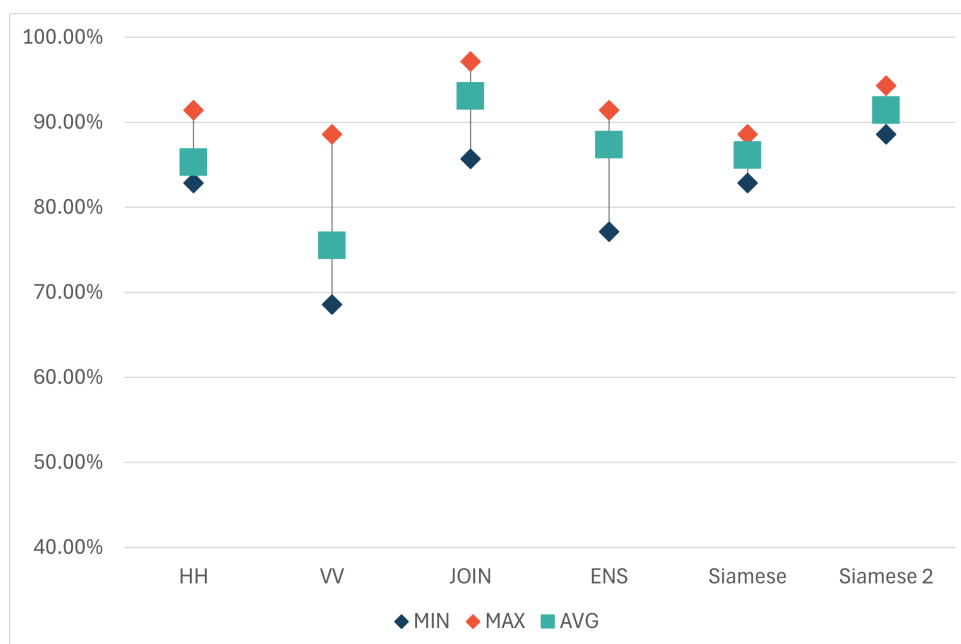
modela. Težine za dio mreže koji se trenira matricama snimljenim horizontalnom polarizacijom inicijalizirane su težinama horizontalnog modela, a za dio koji se trenira matricama vertikalne polarizacije težinama vertikalnog modela. Težine su se nanovo postavljale za svaku iteraciju kako bi odgovarale primjerima kojima su trenirani horizontalni i vertikalni model. Taj model je u nastavku označen kao 'Sijamski 2'. Takav je pristup s predtreniranim modelom povećao prosječnu točnost s 86.12% na 91.43%, ali je i složeniji od svih dosadašnjih. Zanimljivo je da čak ni takav u prosjeku nije ostvario bolje rezultate od JOIN pristupa koji je puno jednostavniji. Razlog tome može biti prije spomenuti mali skup podataka koji više odgovara jednostavnijim modelima, ali i činjenica da se zajednički vektor prije potpuno povezanog sloja kreira spajanjem dva vektora iste veličine što implicira da je utjecaj obje polarizacije na konačnu točnost jednaka što, ranije smo pokazali, nije tako jer je horizontalni model u gotovo svim iteracijama davao točnije rezultate. Tako je u 10 iteracija ovaj model u odnosu na JOIN pristup bio bolji u tri, jednako točan u jednoj, a lošiji u šest. Dodatni argument za korištenje JOIN pristupa je i jednostavnost mreže. U tablici 5.11 (i na slici 5.12) prikazani su rezultati svih modela druge iteracije u kojoj su, uz 'HH', 'VV' i 'JOIN', testirani zajednički model 'ENS' te 'Sijamski' i 'Sijamski 2'.

**Tablica 5.11:** Najmanja, srednja i najveća postignuta točnost kroz 10 iteracija provedenih na modelima treniranim skupovima HH, VV, zajednički (ENS), JOIN, Sijamski bez inicijalizacije i Sijamski 2 u kojem je korištena inicijalizacija težina.

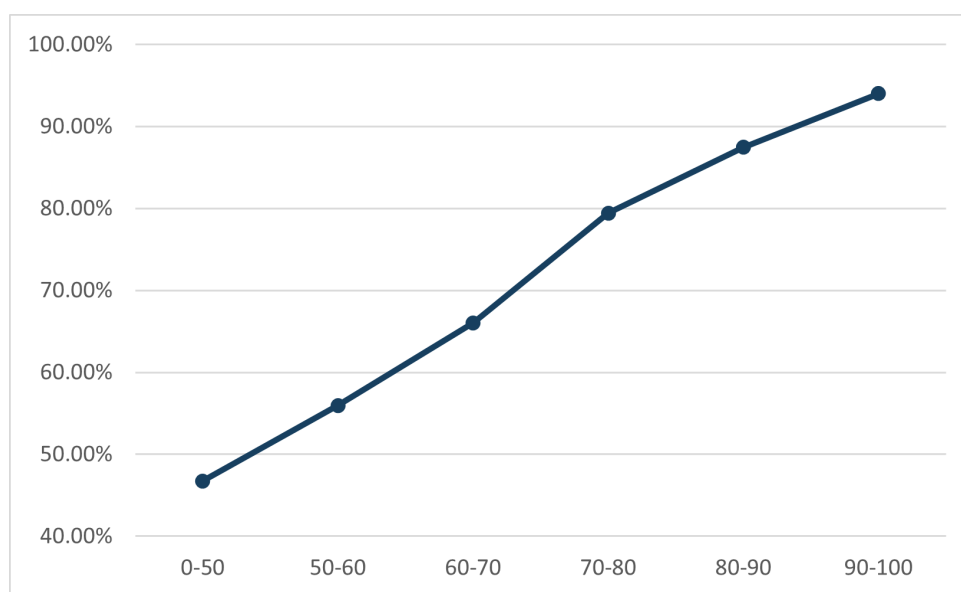
	HH	VV	JOIN	ENS	Sijamski	Sijamski 2
MIN	82.86	68.57	85.71	77.14	82.86	<b>88.57</b>
MAX	91.43	88.57	<b>97.14</b>	91.43	88.57	94.29
AVG	85.31	75.51	<b>93.06</b>	87.34	86.12	91.43

### 5.4.3 Analiza funkcije Softmax

Rezultati svih modela kroz 20 iteracija još su jednom omogućili evaluaciju funkcije softmax kao funkcije kojom se određuje sigurnost modela. Opet su izlazne vjerojatnosti svih 4200 primjera grupirane u grupe 0-50%, 50%-60%, 60%-70%, 70%-80%, 80%-90% i 90%-100% i uspoređeni su broj točnih predviđanja s ukupnim brojem predviđanja svake grupe. Rezultati prikazani u obliku tablice 5.12 i grafa na slici 5.13. potvrdili su da točnost primjera svake grupe upada u tu grupu čime je potvrđena i prije navedena teza o korištenju funkcije softmax za određivanje sigurnosti modela. Zanimljivo je da su primjeri koji upadaju u grupe iznad 80 % zapravo ispravno klasificirani u prosjeku u 92.39% što znači da se taj prag može koristiti u određivanju koje primjere treba snimiti s više pozicija u pristupu povratnom vezom.



**Slika 5.12:** Najmanja, najveća i prosječna točnost modela HH, VV, JOIN, ENS, Sijamski i Sijamski 2 kroz drugi set od 10 iteracija testiranja [94].



**Slika 5.13:** Postotak ispravno klasificiranih u svakoj grupi korištenoj za evaluaciju funkcije Softmax.

#### 5.4.4 Utjecaj polarizacije na klasifikaciju određenog objekta

Ranije je opisano kako su modeli temeljeni na matricama horizontalne polarizacije postizali bolje rezultate u odnosu na one temeljene na matricama vertikalne. U prvih je 10 iteracija razlika prosjeka ta dva modela bila oko 14 postotnih bodova (75% naprema 89%) dok se u drugoj zabilježila nešto niža razlika od 10 postotnih bodova (75% naprema 85%). U samo jednoj iteraciji je zabilježena veća točnost vertikalnog modela.

Kako bi se provjerilo kako polarizacija utječe na konkretne testne objekte korištene prilikom

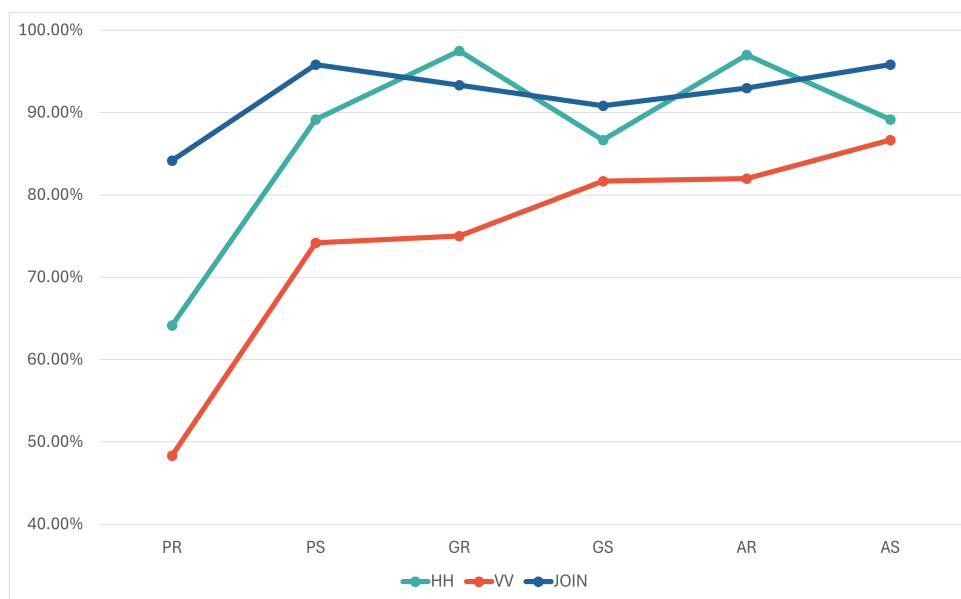
**Tablica 5.12:** Broj ispravno klasificiranih primjera, ukupno primjera i njihov omjer svake grupe za evaluaciju funkcije Softmax kao metode za određivanje sigurnosti modela.

	<b>Ispravno klasificirano</b>	<b>Ukupno primjera</b>	<b>Postotak</b>
0-50	137	293	46.76%
50-60	202	361	55.96%
60-70	245	371	66.04%
70-80	371	467	79.44%
80-90	593	678	87.46%
90-100	1906	2027	94.03%

snimanja provedena je analiza na svih 20 iteracija horizontalnog i vertikalnog modela. Za svaki korišteni objekt uspoređeno je u koliko testnih primjera je ispravno klasificiran i, ako nije, s kojim objektom je najčešće zamijenjen. Skup podataka detaljnije je opisan u poglavlju 4.5, ali može se ponoviti kako su korišteni objekti bili od tri različita materijala (metal, staklo i plastika) i dva različita oblika (valjak i kvadar).

Kod horizontalne polarizacije svi su objekti osim plastičnog valjka (64.17%) klasificirani točnošću iznad 85%. Plastični je valjak u 33% slučajeva klasificiran kao plastični kvadar, dok je broj zamjena ostalim klasama zanemariv. Plastični kvadar je ispravno klasificiran u 89.17% slučajeva, a nije primijećena klasa koja se istaknula u odnosu na druge kad to nije bilo tako. S druge strane, iako su u 86.67%, odnosno 89.17% primjera stakleni i metalni kvadar ispravno klasificirani, kad nisu - stakleni je kvadar najčešće zamijenjen metalnim, a metalni kvadar staklenim. Što se tiče preostala dva objekta (stakleni i metalni valjak), oni su ispravno klasificirani u više od 97% slučajeva kad su snimani horizontalnom polarizacijom. Kod vertikalne je polarizacije samo metalni kvadar ispravno klasificiran u više od 85% slučajeva, dok je najniža točnost, kao i kod horizontalne, zabilježena za plastični valjak (48.33%). Ta je klasa u 45% slučajeva zamijenjena plastičnim kvadrom, a plastični je kvadar najčešće zamijenjen staklenim valjkom. S druge strane, kad su metalni i stakleni objekti pogrešno klasificirani, to je u većini slučajeva bilo zbog materijala, a ne oblika. Točne i predviđene klase kroz 20 iteracija za horizontalnu polarizaciju dane su u tablici 5.13, a za vertikalnu u tablici 5.14.

Na isti je način analiziran i JOIN pristup kod kojeg su svi objekti klasificirani točnošću iznad 90% osim plastičnog valjka (84.17%) koji je, opet, najčešće zamijenjen plastičnim kvadrom. Rezultati tog modela za sve klase kroz svih 20 iteracija dani su u tablici 5.15. Usporedba sva tri pristupa po klasama dana je slikom 5.14.



**Slika 5.14:** Usporedba prosječne točnosti po klasama modela treniranog podacima dobivenim horizontalnom polarizacijom (HH), vertikalnom polarizacijom (VV) i JOIN pristupom [94].

**Tablica 5.13:** Matrica zabune modela treniranog primjerima snimljenim horizontalnom polarizacijom (HH). Klase su označene na sljedeći način: PR - plastični valjak, PS - plastični kvadar, GR - stakleni valjak, GS - stakleni kvadar, AR - metalni valjak, AS - metalni kvadar.

		Predviđena klasa						Ukupno	Točnost
		PR	PS	GR	GS	AR	AS		
Točna klasa	PR	77	40	0	1	2	0	120	64.17%
	PS	1	107	1	4	3	4	120	89.17%
	GR	0	0	117	2	1	0	120	97.50%
	GS	0	2	0	104	0	14	120	86.67%
	AR	0	0	3	0	97	0	100	97.00%
	AS	0	0	0	13	0	107	120	89.17%
Ukupno		1	42	4	20	6	18	700	



**Tablica 5.14:** Matrica zabune modela treniranog primjerima snimljenim vertikalnom polarizacijom (VV). Klase su označene na sljedeći način: PR - plastični valjak, PS - plastični kvadar, GR - stakleni valjak, GS - stakleni kvadar, AR - metalni valjak, AS - metalni kvadar.

		Predviđena klasa						Ukupno	Točnost
		PR	PS	GR	GS	AR	AS		
Točna klasa	PR	58	54	0	8	0	0	120	48.33%
	PS	7	89	13	2	9	0	120	74.17%
	GR	8	5	90	7	10	0	120	75.00%
	GS	2	6	2	98	1	0	120	81.67%
	AR	0	4	12	2	82	11	100	82.00%
	AS	0	0	0	14	2	104	120	86.67%
Ukupno		17	69	27	33	22	18	700	

**Tablica 5.15:** Matrica zabune modela treniranog primjerima JOIN skupa. Klase su označene na sljedeći način: PR - plastični valjak, PS - plastični kvadar, GR - stakleni valjak, GS - stakleni kvadar, AR - metalni valjak, AS - metalni kvadar.

		Predviđena klasa						Ukupno	Točnost
		PR	PS	GR	GS	AR	AS		
Točna klasa	PR	101	18	0	1	0	0	120	84.17%
	PS	0	115	0	0	2	3	120	95.83%
	GR	3	1	112	0	4	0	120	93.33%
	GS	0	4	2	109	0	5	120	90.83%
	AR	2	5	0	0	93	0	100	93.00%
	AS	0	0	0	4	1	115	120	95.83%
Ukupno		5	28	2	5	7	8	700	

## Poglavlje 6

# Sustav za snimanje SAR-om i integriranom klasifikacijom objekata

Predložena je implementacija energetski učinkovitog sustava koji objedinjuje radar sa sintetičkom aperturom i model za klasifikaciju objekata na radarskim podacima. Sustav se temelji na ideji korištenja mikroracunala kojim se omogućava komunikacija između te dvije komponente. Komunikacijom se, pokazano je, može smanjiti potrošnja energije za vrijeme snimanja radarske slike zadržavajući visoku točnost klasifikacije. U našem slučaju sustav je izveden pomoću mikroracunala Raspberry Pi. Implementacija radarske komponente sustava opisana je u poglavlju 2.3. Sustav GBSAR-Pi zasniva se na kontroliranom kretanju platforme na kojoj se nalazi FMCW modul koji rezultatne signale snimljene s određenog broja pozicija na vodilici šalje mikroracunalu. Skup svih snimljenih signala može se obraditi algoritmom za rekonstrukciju radarske slike. Implementacija algoritma Omega-K kojim to radi predloženi sustav opisana je u poglavlju 2.3.2. Kao što je navedeno u poglavlju 5.1, klasifikacija se može raditi na dva načina: na rekonstruiranim slikama ili sirovim radarskim podacima. Druga opcija, pokazano je, daje nešto bolje rezultate, ali je podatke zapisane u takvom obliku teže interpretirati. Zato se predlaže obrada rekonstrukcijskim algoritmom onih primjera koje je potrebno dodatno analizirati.

Za klasifikaciju objekata na sirovim podacima predložen je model temeljen na arhitekturi ResNet-18. Prilagodba te arhitekture podacima u obliku matrice u kojoj je jedna dimenzija puno veća od druge zasniva se na ideji uklanjanja koraka poduzorkovanja za manju dimenziju. Ta je prilagodba detaljnije opisana u poglavlju 5.1.2, a takav je model u radu nazvan 'Modificirani ResNet-18'. Predloženim modelom analizirano je nekoliko načina prikupljanja podataka razvijenim sustavom, a svaki od njih rezultirao je zasebnim skupom podataka. Svojstva kreiranih skupova podataka dana su u poglavlju 4. Analize su potvrdile utjecaj parametara radara na količinu informacije snimljenih podataka smanjenjem točnosti klasifikacije u slučaju uže širine frekvencijskog pojasa i duljeg koraka GBSAR-a što je objašnjeno u 5.2. S druge strane, pokazano je i da se u slučaju duljeg koraka uz dovoljno veliki skup podataka mogu dobiti visoke

točnosti klasifikacije pa je predložen način snimanja ovog sustava u dvije iteracije. U prvoj su objekti promatrani s manjeg broja pozicija (dulji korak GBSAR-a) i na tim se podacima radi inicijalna klasifikacija, a u drugoj su objekti za koje je vjerojatnost modela ispod praga snimani s dodatnih pozicija nakon čega slijedi završna klasifikacija. Ta je komunikacija između modela i radara u radu nazvana 'povratna veza' i rezultati tog načina snimanja opisani su u poglavlju 5.3. Budući da je moguće mijenjati polarizaciju na radarskoj komponenti razvijenog sustava, testirana je opcija snimanja iste scene s više polarizacija kako bi se dobila dodatna informacija koja se može iskoristiti za povećanje točnosti klasifikacije. Analiza nekoliko raličitih kombinacija dviju matrica (po jedna za horizontalnu i vertikalnu polarizaciju) istaknula je pristup u kojem su dvije matrice slijepljene jedna za drugu (JOIN) kao najbolji. Svi su pristupi predstavljeni u poglavlju 5.4.

Na kraju, predložen je sustav za klasifikaciju objekata na radarskim podacima i sljedeći način rada: Podaci se snimaju s manjeg broja pozicija na vodilici koristeći dvije polarizacije, a rezultatne matrice se u obradi povezuju JOIN pristupom. Nakon inicijalne klasifikacije modelom prilagođenim za rad sa sirovim podacima, dijelovi scene u kojima se nalazi objekt za koji model nije sa sigurnošću mogao odrediti klasu promatraju se s dodatnih pozicija. Sigurnost modela određuje se funkcijom softmax opisanom u poglavlju 3.1.2, a za prag se predlaže 80% (objašnjeno u poglavlju 5.4.3). U dodatnom snimanju spomenutih dijelova scene opet se koriste dvije polarizacije i JOIN pristup povezivanja matrica. Važno je napomenuti kako se signale dobivene dodatnim snimanjem prije klasifikacije spaja s onima inicijalnog snimanja. To je moguće jer se pozicije FMCW modula za vrijeme dodatnog snimanja ne preklapaju s onima inicijalnog. Nakon završnog snimanja još se jednom modelom 'Modificirani ResNet-18' klasificiraju objekti koji su u inicijalnoj klasifikaciji bili ispod praga. Rezultati završne klasifikacije konačni su rezultati sustava GBSAR-Pi.

# Poglavlje 7

## Zaključak

Modeli dubokog učenja trenirani slikama snimljenim radarom sa sintetičkom aperturom (eng. *Synthetic Aperture Radar*, SAR) koji kretanjem senzora pospješuje rezoluciju imaju sve širu primjenu. Radarskim slikama satelitskog SAR-a u kojem se senzor na satelitu kreće po orbiti, takvi modeli najčešće detektiraju vozila, klasificiraju vegetaciju i prate promjene mora i oceana. Ipak, snimanje iz svemira ima svoje nedostatke. Osim što se određene frekvencije ne mogu koristiti zbog molekula vode u atmosferi, vrijeme ponovnog snimanja istog područja kod satelita je često iskazano u tjednima. Spuštanjem SAR koncepta na tlo (eng. *Ground Based SAR*, GBSAR) povećava se fleksibilnost prilikom odabira područja interesa, smanjuje vrijeme ponavljanja snimanja na nekoliko minuta i omogućava korištenje viših frekvencija čime se dodatno poboljšava rezolucija radara. GBSAR stoga može milimetarskom preciznošću nadzirati infrastrukturne objekte i reljefne promjene. Novi niz mogućnosti otvara se objedinjavanjem modela dubokog učenja i radara sa sintetičkom aperturom u jedan sustav čime se omogućava dinamička prilagodba parametara radara problemu koji model rješava, a upravo je to glavna ideja ovog rada.

U radu je predstavljen sustav GBSAR-Pi koji se sastoji od GBSAR-a i modela za klasifikaciju objekata. Naziv je dobio po mikroročunalu Raspberry Pi oko kojeg je razvijen. Testirano je nekoliko arhitektura modela dubokog učenja, a najboljom se pokazala arhitektura prilagođena uređajima manje procesorske snage - ResNet-18. Radarsko snimanje ostvareno je kretanjem senzora po vodilici stani-kreni načinom kojim se poništava utjecaj pomicanja senzora za vrijeme mjerenja na određenoj poziciji. Korišteni senzor je FMCW (od eng. *Frequency Modulated Continuous Wave*) modul IVS-362 s centralnom frekvencijom od 24 GHz i programabilnom širinom pojasa najvišeg iznosa 1.3 GHz. Nakon snimanja sirovi podaci obrađeni su rekonstrukcijskim algoritmom Omega-K kojim se generira radarska slika promatranog prostora.

Sustavom GBSAR-Pi kreirano je i javno objavljeno pet skupova radarskih podataka koji se međusobno razlikuju u promatranj sceni i objektima, korištenim radarskim parametrima prilikom snimanja i generiranim rekonstruiranim slikama. Skup koji sadrži sirove i rekons-

truirane radarske podatke omogućio je usporedbu pristupa klasifikaciji u dva smjera: pomoću niza sirovih signala snimljenih sustavom GBSAR-Pi i pomoću radarskih slika rekonstruiranih tim signalima. Pokazalo se da algoritam za rekonstrukciju slike zbog koraka aproksimacije i interpolacije narušava integritet podataka i snižava rezultat klasifikacije, a posebno je zanimljivo da nije zabilježen nijedan testni primjer u kojem je model temeljen na rekonstruiranim slikama ispravno odredio klasu, a onaj temeljen na sirovim podacima pogrešno. Taj je rezultat i očekivan budući da algoritam za rekonstrukciju, iako olakšava ljudsku interpretaciju snimljene scene, ne dodaje informaciju koja bi mogla poboljšati klasifikaciju objekata na njoj. Zato su sljedeći modeli trenirani sirovim radarskim podacima. Klasifikacija tim pristupom dodatno je poboljšana modifikacijom arhitekture modela. Zbog specifičnosti izlazne matrice snimljenih signala u kojoj je jedna dimenzija puno veća od druge, izmijenjena je arhitektura ResNet-18 uklanjanjem koraka poduzorkovanja za manju dimenziju kako bi se modelu omogućilo korištenje informacije koju daje korelacija između mjerenja senzora s različitim pozicija.

Količina informacije koju izlazna matrica nosi ovisi o parametrima radara. Ta se količina na rekonstruiranoj slici očituje u rezoluciji, a u oba prije navedena pristupa klasifikaciji u točnosti modela. Kako bi se analizirao efekt promjene parametara, uspoređeni su rezultati klasifikacijskih modela trenirani podacima snimljenim sustavom GBSAR-Pi koristeći različite širine frekvencijskog pojasa i udaljenosti između dvije uzastopne pozicije senzora s kojih se obavlja FMCW mjerenje. Promjenom prvog parametra utjecalo se na signale snimljene na svakoj poziciji, a drugim na broj tih pozicija što znači da je integritet podataka narušen u oba smjera - po udaljenosti i po smjeru kretanja senzora. Model testiran na podacima snimljenim kombinacijom u kojoj je širina frekvencijskog pojasa najveća i u kojoj je udaljenost između dviju pozicija najmanja ostvario je najbolje rezultate što je očekivano budući da je količina informacije koju matrica snimljena tom kombinacijom ima - najveća. Ipak, pokazalo se kako je, uz dovoljno veliki skup podataka, moguće zadržati visoku razinu točnosti klasifikacije i kad se koristi manji broj pozicija s kojih se obavlja mjerenje senzora što je zanimljivo iz perspektive energetske učinkovitosti cijelog sustava.

Implementacija sustava GBSAR-Pi na jednom mikrorračunalu omogućila je izravnu komunikaciju između modela dubokog učenja i radara, odnosno omogućila je dinamičko postavljanje radarskih parametara temeljem rezultata klasifikacije modela. Budući da na potrošnju energije najviše utječe broj pozicija s kojih senzor obavlja mjerenje, komunikacijom ta dva podsustava ostvaren je balans između točnosti modela i broja pozicija čime sustav postaje energetski učinkovitiji. U radu je opisan predloženi način snimanja na primjeru scene s tri objekta međusobno jednako udaljena. Inicijalno snimanje provodi se s manjeg broja pozicija nakon čega klasifikacijski model određuje koje dijelove scene treba detaljnije snimiti s dodatnih pozicija u završnom snimanju. Na taj je način sustav u jednom od opisanih primjera scenu promatrao sa 62.5% manje pozicija nego u regularnom snimanju, a postigao je usporedivu razinu točnosti

- iznad 90%. Dijelovi scene koji se trebaju dodatno snimiti određuju se pomoću vjerojatnosti modela za najizgledniju klasu koja se računa funkcijom softmax. Na testnim primjerima kroz nekoliko iteracija pokazano je da se takvim pristupom može ispravno odrediti vjerojatnost.

Problem smanjenja broja koraka GBSAR-a je spomenuta manja količina informacija koju izlazna matrica nosi. Ta se količina može povećati dodavanjem novih izvora informacija. U radu je opisana mogućnost korištenja različitih polarizacija elektromagnetskih valova koje senzor odašilje za vrijeme snimanja kako bi se ista scena snimila na više načina, odnosno pomoću više polarizacija. Analizirane su razlike u podacima snimljenim horizontalnom (HH) i vertikalnom (VV) polarizacijom i uspoređeni rezultati modela trenirani na njima. Budući da modeli nisu bili jednako točni u klasifikaciji testnih objekata, može se zaključiti da podaci snimljeni različitim polarizacijama nose različite informacije. U sljedećem su koraku analizirane metode povezivanja dviju izlaznih matrica (po jedna za svaku promatranu polarizaciju) od kojih se posebno istaknula metoda koja spaja jednu matricu na kraj druge (JOIN metoda). Modeli trenirani podacima generiranim tom metodom u odnosu na modele trenirane podacima jedne polarizacije ostvarili su u prosjeku 2, odnosno 8 postotnih bodova višu točnost kroz dva seta od 10 iteracija testiranja. Problem takvog pristupa je što model umjesto dvije različite izlazne matrice na ulaz prima jednu zajedničku. Zato je implementiran Sijamski model čija arhitektura omogućava primanje dva zasebna ulaza za koje postavlja težine na isti način kao u modificiranom ResNet modelu, ali prije izlaza spaja vektore značajki u jedan zajednički koji nakon potpuno povezanog sloja daje jedinstveni izlaz. Sijamski model s incijaliziranim težinama kroz 10 iteracija postizao je usporedivu točnost kao i JOIN model.

Predloženi sustav za radarsko snimanje i klasifikaciju objekata GBSAR-Pi koristi princip radara sa sintetičkom aperturom u dva navrata između kojih povratnom vezom postavlja radarske parametre. Tijekom snimanja odašilje EM valove horizontalne i vertikalne polarizacije i dvije izlazne matrice povezuje JOIN metodom. Klasifikacijski model temeljen je na ResNet-18 arhitekturi prilagođenoj radu sa sirovim radarskim podacima.

# Literatura

- [1] Turing, A. M., “Computing Machinery and Intelligence”, *Mind*, Vol. 59, No. October, 1950, str. 433–60.
- [2] Wiley, C. A., “Synthetic Aperture Radars”, *IEEE Transactions on Aerospace and Electronic Systems*, No. 3, 1985, str. 440–443.
- [3] He, K., Zhang, X., Ren, S., Sun, J., “Deep Residual Learning for Image Recognition”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, str. 770–778.
- [4] Hochreiter, S., Schmidhuber, J., “Long Short-Term Memory”, *Neural computation*, Vol. 9, No. 8, 1997, str. 1735–1780.
- [5] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”, *CoRR*, Vol. abs/1704.04861, 2017, dostupno na: <http://arxiv.org/abs/1704.04861>
- [6] Fukushima, K., “Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position”, *Biological cybernetics*, Vol. 36, No. 4, 1980, str. 193–202.
- [7] Medsker, L. R., Jain, L., “Recurrent Neural Networks”, *Design and Applications*, Vol. 5, No. 64-67, 2001, str. 2.
- [8] Elman, J. L., “Finding Structure in Time”, *Cognitive Science*, Vol. 14, No. 2, 1990, str. 179-211, dostupno na: <https://www.sciencedirect.com/science/article/pii/036402139090002E>
- [9] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N. *et al.*, *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. Red Hook, NY, USA: Curran Associates Inc., 2019.

- [10]Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. *et al.*, “TensorFlow: A System for Large-Scale Machine Learning”, in 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16), 2016, str. 265–283.
- [11]Kingma, D. P., Ba, J., “Adam: A Method for Stochastic Optimization”, arXiv preprint arXiv:1412.6980, 2014.
- [12]Duchi, J., Hazan, E., Singer, Y., “Adaptive Subgradient Methods for Online Learning and Stochastic Optimization”, Journal of machine learning research, Vol. 12, No. 7, 2011.
- [13]Brown, W. M., “Synthetic Aperture Radar”, IEEE Transactions on Aerospace and Electronic Systems, No. 2, 1967, str. 217–229.
- [14]Cumming, I., Bennett, J., “Digital Processing of Seasat SAR data”, in ICASSP ’79. IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 4, 1979, str. 710-718.
- [15]Raney, R., Runge, H., Bamler, R., Cumming, I., Wong, F., “Precision SAR processing using chirp scaling”, IEEE Transactions on Geoscience and Remote Sensing, Vol. 32, No. 4, 1994, str. 786-799.
- [16]Cafforio, C., Prati, C., Rocca, F., “SAR Data Focusing Using Seismic Migration Techniques”, IEEE transactions on aerospace and electronic systems, Vol. 27, No. 2, 1991, str. 194–207.
- [17]Andersson, L.-E., “On the Determination of a Function from Spherical Averages”, SIAM Journal on Mathematical Analysis, Vol. 19, No. 1, 1988, str. 214-232, dostupno na: <https://doi.org/10.1137/0519016>
- [18]Ulander, L., Hellsten, H., Stenstrom, G., “Synthetic-Aperture Radar Processing Using Fast Factorized Back-Projection”, IEEE Transactions on Aerospace and Electronic Systems, Vol. 39, No. 3, 2003, str. 760-776.
- [19]Vu, V. T., Sjogren, T. K., Pettersson, M. I., “A Comparison Between Fast Factorized Back-projection and Frequency-Domain Algorithms in UWB Lowfrequency SAR”, in IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium, Vol. 4, 2008, str. IV - 1284-IV - 1287.
- [20]Potsis, A., Reigber, A., Alivizatos, E., Moreira, A., Uzunoglou, N. K., “Comparison of Chirp Scaling and Wavenumber Domain Algorithms for Airborne Low-Frequency SAR Data Processing”, in SAR Image Analysis, Modeling, and Techniques V, Vol. 4883. SPIE, 2003, str. 11–19.



- [21] Bamler, R., "A Comparison of Range-Doppler and Wavenumber Domain SAR Focusing Algorithms", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, No. 4, 1992, str. 706-713.
- [22] Ribalta, A., "Time-Domain Reconstruction Algorithms for FMCW-SAR", *IEEE Geoscience and Remote Sensing Letters*, Vol. 8, No. 3, 2011, str. 396-400.
- [23] Cumming, I., Neo, Y., Wong, F., "Interpretations of the Omega-K Algorithm and Comparisons with other Algorithms", in *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)*, Vol. 3, 2003, str. 1455-1458.
- [24] Stolt, R. H., "Migration by Fourier Transform", *Geophysics*, Vol. 43, No. 1, 1978, str. 23-48.
- [25] Guo, S., Dong, X., "Modified Omega-K Algorithm for Ground-Based FMCW SAR Imaging", in *2016 IEEE 13th International Conference on Signal Processing (ICSP)*, 2016, str. 1647-1650.
- [26] Meta, A., Hoozeboom, P., Ligthart, L. P., "Signal Processing for FMCW SAR", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, No. 11, 2007, str. 3519-3532.
- [27] Rudolf, H., Leva, D., Tarchi, D., Sieber, A., "A Mobile and Versatile SAR System", in *IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No.99CH36293)*, Vol. 1, 1999, str. 592-594 vol.1.
- [28] Martinez-Vazquez, A., Fortuny-Guasch, J., "A GB-SAR Processor for Snow Avalanche Identification", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 11, 2008, str. 3948-3956.
- [29] Martinez-Vazquez, A., Fortuny-Guasch, J., "3 Winters of Snow Monitoring With the LISA Instrument in the Sion Valley (CH): 2003-2006", *Off. Official Publ. Eur. Commun., Luxembourg, Belgium, Tech. Rep. EUR*, Vol. 22533.
- [30] Bennett, J., Morrison, K., Race, A., Cookmartin, G., Quegan, S., "The UK NERC Fully Portable Polarimetric Ground-Based Synthetic Aperture Radar (GB-SAR)", in *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No.00CH37120)*, Vol. 5, 2000, str. 2313-2315 vol.5.
- [31] Zhou, Z.-S., Boerner, W.-M., Sato, M., "Development of a Ground-Based Polarimetric Broadband SAR System for Noninvasive Ground-Truth Validation in Vegetation Monito-

- ring”, IEEE Transactions on Geoscience and Remote Sensing, Vol. 42, No. 9, 2004, str. 1803-1810.
- [32]Florentino, A., Charapaqui, S., De La Jara, C., Milla, M., “Implementation of a Ground Based Synthetic Aperture Radar (GB-SAR) for Landslide Monitoring: System Description and Preliminary Results”, in 2016 IEEE XXIII International Congress on Electronics, Electrical Engineering and Computing (INTERCON), 2016, str. 1-6.
- [33]Feng, W., Friedt, J.-M., Wan, P., “SDR-Implemented Ground-Based Interferometric Radar for Displacement Measurement”, IEEE Transactions on Instrumentation and Measurement, Vol. 70, 2021, str. 1-18.
- [34]Amézaga, A., López-Martínez, C., Jové, R., “A Multi-Frequency SDR-Based GBSAR: System Overview and First Results”, Remote Sensing, Vol. 13, No. 9, 2021, dostupno na: <https://www.mdpi.com/2072-4292/13/9/1613>
- [35]Li, J., Zheng, X., Stoica, P., “Mimo SAR Imaging: Signal Synthesis and Receiver Design”, in 2007 2nd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2007, str. 89-92.
- [36]Li, J., Stoica, P., MIMO Radar Signal Processing. John Wiley & Sons, 2008.
- [37]Miccinesi, L., Consumi, T., Beni, A., Pieraccini, M., “W-band MIMO GB-SAR for Bridge Testing/Monitoring”, Electronics, Vol. 10, No. 18, 2021, dostupno na: <https://www.mdpi.com/2079-9292/10/18/2261>
- [38]Tarchi, D., Oliveri, F., Sammartino, P. F., “MIMO Radar and Ground-Based SAR Imaging Systems: Equivalent Approaches for Remote Sensing”, IEEE Transactions on Geoscience and Remote Sensing, Vol. 51, No. 1, 2013, str. 425-435.
- [39]U.S. Air Force, “MSTAR Overview”, dostupno na: <https://www.sdms.afrl.af.mil/index.php?collection=mstar> , pristupano 11.1.2024.
- [40]Keydel, E. R., Lee, S. W., Moore, J. T., “MSTAR Extended Operating Conditions: a Tutorial”, in Algorithms for Synthetic Aperture Radar Imagery III, Zelnio, E. G., Douglass, R. J., (ur.), Vol. 2757, International Society for Optics and Photonics. SPIE, 1996, str. 228 – 242, dostupno na: <https://doi.org/10.1117/12.242059>
- [41]Yang, Y., Qiu, Y., Lu, C., “Automatic Target Classification - Experiments on the MSTAR SAR Images”, in Sixth International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Network, 2005, str. 2-7.

- [42]Cover, T., Hart, P., “Nearest Neighbor Pattern Classification”, IEEE Transactions on Information Theory, Vol. 13, No. 1, 1967, str. 21-27.
- [43]Wang, H., Chen, S., Xu, F., Jin, Y.-Q., “Application of Deep-Learning Algorithms to MSTAR Data”, in 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, str. 3743-3745.
- [44]Blasch, E., Majumder, U., Zelnio, E., Velten, V., “Review of Recent Advances in AI/ML Using the MSTAR Data”, in Algorithms for Synthetic Aperture Radar Imagery XXVII, Zelnio, E., Garber, F. D., (ur.), Vol. 11393, International Society for Optics and Photonics. SPIE, 2020, str. 113930C, dostupno na: <https://doi.org/10.1117/12.2559035>
- [45]Profeta, A., Rodriguez, A., Clouse, H. S., “Convolutional Neural Networks for Synthetic Aperture Radar Classification”, in Algorithms for synthetic aperture radar imagery XXIII, Vol. 9843. SPIE, 2016, str. 185–194.
- [46]Majumder, U., Christiansen, E., Wu, Q., Inkawhich, N., Blasch, E., Nehrbass, J., “High-Performance Computing for Automatic Target Recognition in Synthetic Aperture Radar Imagery”, in Cyber Sensing 2017, Ternovskiy, I. V., Chin, P., (ur.), Vol. 10185, International Society for Optics and Photonics. SPIE, 2017, str. 1018508, dostupno na: <https://doi.org/10.1117/12.2263218>
- [47]Srinivas, U., Monga, V., Raj, R. G., “SAR Automatic Target Recognition Using Discriminative Graphical Models”, IEEE Transactions on Aerospace and Electronic Systems, Vol. 50, No. 1, 2014, str. 591-606.
- [48]Zhang, F., Hu, C., Yin, Q., Li, W., Li, H.-C., Hong, W., “Multi-Aspect-Aware Bidirectional LSTM Networks for Synthetic Aperture Radar Target Recognition”, IEEE Access, Vol. 5, 2017, str. 26 880-26 891.
- [49]Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., Hossard, L., “Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France”, Remote Sensing, Vol. 10, No. 8, 2018, dostupno na: <https://www.mdpi.com/2072-4292/10/8/1217>
- [50]Khalili, M. A., Guerriero, L., Pouralizadeh, M., Calcaterra, D., Di Martire, D., “Monitoring and Prediction of Landslide-Related Deformation Based on the GCN-LSTM Algorithm and SAR Imagery”, Natural Hazards, Vol. 119, No. 1, 2023, str. 39–68.
- [51]Grohnfeldt, C., Schmitt, M., Zhu, X., “A Conditional Generative Adversarial Network to Fuse Sar And Multispectral Optical Data For Cloud Removal From Sentinel-2 Images”,

- in IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018, str. 1726-1729.
- [52]Ebel, P., Schmitt, M., Zhu, X. X., “Cloud Removal in Unpaired Sentinel-2 Imagery Using Cycle-Consistent GAN and SAR-Optical Data Fusion”, in IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, 2020, str. 2065-2068.
- [53]Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., “Generative adversarial nets”, in Advances in Neural Information Processing Systems, 2014, str. 2672–2680.
- [54]Zhao, Y., Celik, T., Liu, N., Li, H.-C., “A Comparative Analysis of GAN-Based Methods for SAR-to-Optical Image Translation”, IEEE Geoscience and Remote Sensing Letters, Vol. 19, 2022, str. 1-5.
- [55]Han, J., Yang, H., Liu, Y., Lu, Z., Zeng, K., Jiao, R., “A Deep Learning Application for Deformation Prediction from Ground-Based InSAR”, Remote Sensing, Vol. 14, No. 20, 2022, dostupno na: <https://www.mdpi.com/2072-4292/14/20/5067>
- [56]Ma, Z., Mei, G., Prezioso, E., Zhang, Z., Xu, N., “A Deep Learning Approach Using Graph Convolutional Networks for Slope Deformation Prediction Based on Time-Series Displacement Data”, Neural Computing and Applications, Vol. 33, No. 21, 2021, str. 14 441–14 457.
- [57]Du, S., Feng, G., Wang, J., Feng, S., Malekian, R., Li, Z., “A New Machine-Learning Prediction Model for Slope Deformation of an Open-Pit Mine: An Evaluation of Field Data”, Energies, Vol. 12, No. 7, 2019, dostupno na: <https://www.mdpi.com/1996-1073/12/7/1288>
- [58]Kang, M.-K., Kim, K.-E., Cho, S.-J., Lee, H., Lee, J.-H., “Wishart Supervised Classification Results of C-band Polarimetric GB-SAR Image Data”, in 2011 IEEE International Geoscience and Remote Sensing Symposium, 2011, str. 459-462.
- [59]European Space Agency, “Copernicus Dataspace”, dostupno na: <https://dataspace.copernicus.eu/> , pristupano 12.1.2024.
- [60]National Aeronautics and Space Agency, “NASA EarthData”, dostupno na: <https://search.earthdata.nasa.gov> , pristupano 12.1.2024.
- [61]Tur činović, F., “Ground Based SAR Data for Classification - 9 Objects in the Near Distance”, dostupno na: <https://data.mendeley.com/datasets/p2yhyx7335/1> 2023.

- [62]Tur činović, F., “Ground Based SAR Data Obtained With Different Combinations of Bandwidth and Step Size”, dostupno na: <https://data.mendeley.com/datasets/vpmnxw5nng/1> 2023.
- [63]Tur činović, F., “Ground Based SAR Data of Cardboard box with Aluminium Bottles for Anomaly Detection”, dostupno na: <https://data.mendeley.com/datasets/s8py5wdk4g/1> 2022.
- [64]Tur činović, F., “Near-Distance Raw and Reconstructed Ground Based SAR Data”, dostupno na: <https://data.mendeley.com/datasets/m458grc688/2> 2023.
- [65]Tur činović, F., “Ground Based SAR Data Obtained With Different Polarizations”, dostupno na: <https://data.mendeley.com/datasets/nbc9xpww96/1> 2023.
- [66]Cortes, C., Vapnik, V., “Support-Vector Networks”, *Machine Learning*, Vol. 20, No. 3, 1995, str. 273–297.
- [67]Lee, J., Grunes, M., “Classification of Multi-Look Polarimetric SAR data Based on Complex Wishart Distribution”, in [Proceedings] NTC-92: National Telesystems Conference, 1992, str. 7/21-7/24.
- [68]Chen, H., Zeng, Z., “Deformation Prediction of Landslide Based on Improved Back-Propagation Neural Network”, *Cognitive computation*, Vol. 5, 2013, str. 56–62.
- [69]Jang, J.-S., “ANFIS: Adaptive-Network-Based Fuzzy Inference System”, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 23, No. 3, 1993, str. 665-685.
- [70]Tipping, M., “The Relevance Vector Machine”, *Advances in neural information processing systems*, Vol. 12, 1999.
- [71]Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H., “T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction”, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, No. 9, 2020, str. 3848-3858.
- [72]Gori, M., Monfardini, G., Scarselli, F., “A New Model for Learning in Graph Domains”, in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, Vol. 2. IEEE, 2005, str. 729–734.
- [73]Woodhouse, I. H., *Introduction to Microwave Remote Sensing*. CRC press, 2017.
- [74]Miller, P. D., *Applied Asymptotic Analysis*. American Mathematical Soc., 2006, Vol. 75.
- [75]Yamaguchi, Y., *Polarimetric SAR Imaging: Theory and Applications*. CRC Press, 2020.

- [76]Lee, J.-S., Ainsworth, T. L., “An Overview of Recent Advances in Polarimetric SAR Information Extraction: Algorithms and Applications”, in 2010 IEEE International Geoscience and Remote Sensing Symposium, 2010, str. 851-854.
- [77]Hamasaki, T., Ferro-Famil, L., Pottier, E., Sato, M., “Applications of Polarimetric Interferometric Ground-Based SAR (GB-SAR) System to Environment Monitoring and Disaster Prevention”, in European Radar Conference, 2005. EURAD 2005., 2005, str. 29-32.
- [78]Raspberry Pi Foundation, “Raspberry Pi 4B Specifications”, dostupno na: <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/specifications/> , pristupano 15.1.2024.
- [79]Innovative Radar Sensor Technology, “Radar Sensor IVS-362”, dostupno na: <https://www.innosent.de/en/sensors/ivs-362/> , pristupano 15.1.2024.
- [80]Kačan, M., Turčinović, F., Bojanjac, D., Bosiljevac, M., “Deep Learning Approach for Object Classification on Raw and Reconstructed GBSAR Data”, Remote Sensing, Vol. 14, No. 22, 2022, dostupno na: <https://www.mdpi.com/2072-4292/14/22/5673>
- [81]Kelleher, J. D., Deep Learning. MIT press, 2019.
- [82]Šegvić, S., Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu, “Predavanje Duboko učenje 1”.
- [83]Kačan, M., “Metrička ugrađivanja za reprezentiranje osoba u slikama”, 2019.
- [84]Rumelhart, D. E., Hinton, G. E., Williams, R. J., “Learning Representations by Back-Propagating Errors”, nature, Vol. 323, No. 6088, 1986, str. 533–536.
- [85]Polyak, B., “Some Methods of Speeding up the Convergence of Iteration Methods”, USSR Computational Mathematics and Mathematical Physics, Vol. 4, No. 5, 1964, str. 1-17, dostupno na: <https://www.sciencedirect.com/science/article/pii/0041555364901375>
- [86]Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., “Imagenet: A Large-Scale Hierarchical Image Database”, in 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009, str. 248–255.
- [87]Glorot, X., Bengio, Y., “Understanding the Difficulty of Training Deep Feedforward Neural Networks”, in Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2010, str. 249–256.
- [88]Bengio, Y., Simard, P., Frasconi, P., “Learning Long-Term Dependencies with Gradient Descent is Difficult”, IEEE transactions on neural networks, Vol. 5, No. 2, 1994, str. 157–166.

- [89]Ioffe, S., Szegedy, C., “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift”, in International conference on machine learning. pmlr, 2015, str. 448–456.
- [90]Elsken, T., Metzen, J. H., Hutter, F., “Neural Architecture Search: A Survey”, dostupno na: <https://arxiv.org/abs/1808.05377> 2019.
- [91]Yang, T.-J., Howard, A., Chen, B., Zhang, X., Go, A., Sandler, M., Sze, V., Adam, H., “NetAdapt: Platform-Aware Neural Network Adaptation for Mobile Applications”, dostupno na: <https://arxiv.org/abs/1804.03230> 2018.
- [92]Gu, Q., Zhu, L., Cai, Z., “Evaluation Measures of the Classification Performance of Imbalanced Data Sets”, in Computational Intelligence and Intelligent Systems: 4th International Symposium, ISICA 2009, Huangshi, China, October 23-25, 2009. Proceedings 4. Springer, 2009, str. 461–471.
- [93]Šnajder, J., Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu, “Materijali za Strojno učenje 1: 21. Vrednovanje modela”.
- [94]Tur činović, F., Kačan, M., Bojanjac, D., Bosiljevac, M., Šipuš, Z., “Utilizing Polarization Diversity in GBSAR Data-Based Object Classification”, *Sensors*, Vol. 24, No. 7, 2024, dostupno na: <https://www.mdpi.com/1424-8220/24/7/2305>
- [95]Tur činović, F., Kačan, M., Bojanjac, D., Bosiljevac, M., “Impact of Ground Based SAR Parameters on Radar Data Based Object Classification”, in 2023 24th International Conference on Applied Electromagnetics and Communications (ICECOM), 2023, str. 1-5.
- [96]Tur činović, F., Kačan, M., Bojanjac, D., Bosiljevac, M., “Ground-Based SAR System for Object Classification with Parameter Optimization Based on Deep Learning Feedback Algorithm”, in *Image and Signal Processing for Remote Sensing XXIX*, Bruzzone, L., Bovolo, F., (ur.), Vol. 12733, International Society for Optics and Photonics. SPIE, 2023, str. 127330R, dostupno na: <https://doi.org/10.1117/12.2679852>

# Životopis

Filip Turčinović rođen je u Rijeci 1994. godine. Osnovnu školu završio je u Svetom Petru u Šumi (OŠ Vladimira Nazora), a srednju u Pazinu (Gimnazija i strukovna škola Jurja Dobrile Pazin). 2013. upisao je Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu (FER), 2016. obranio završni rad pod mentorstvom profesora Josipa Knezovića, a 2018. diplomski rad pisan pod mentorstvom profesorice Sonje Grgić. U studenom 2018. godine zaposlio se kao istraživač na Zavodu za komunikacijske i svemirske tehnologije FER-a (ZKIST) na projektu „Razvoj otvorene pametne mreže energetske učinkovite javne LED rasvjete“ pod vodstvom profesora Gordana Šišula. U listopadu 2019. godine upisao je doktorski studij na FER-u na kojem mu je mentor profesor Marko Bosiljevac. Područje istraživanja usko je vezano uz projekt „Pametna platforma za daljinska istraživanja u okolišu i industriji primjenom milimetarskih valova“ na kojem je zaposlen od listopada 2020. S profesorom Darijem Bojanjcem pokrenuo je Institut za promociju znanosti. Ukupno je objavio 16 znanstvenih radova od čega četiri u međunarodnim časopisima. Dobitnik je Rektorove nagrade za timski znanstveni i umjetnički rad 2018. godine.

## Popis objavljenih djela

### Rad u časopisima

1. Turčinović, F., Kačan, M., Bojanjac, D., Bosiljevac, M., Šipuš, Z., "Utilizing Polarization Diversity in GBSAR Data Based Object Classification," *Sensors*, 2024, MDPI.
2. Turčinović, F., Kačan, M., Bojanjac, D., Bosiljevac, M., "Near-Distance Raw and Reconstructed Ground-Based SAR Data," *Data in Brief*, 51, 109620, 2023, Elsevier.
3. Kačan, M., Turčinović, F., Bojanjac, D., Bosiljevac, M., "Deep Learning Approach for Object Classification on Raw and Reconstructed GBSAR Data," *Remote Sensing*, 14(22), 5673, 2022, MDPI.
4. Turčinović, F., Šišul, G., Bosiljevac, M., "LoRaWAN Base Station Improvement for Better Coverage and Capacity," *Journal of Low Power Electronics and Applications*, 12(1), 1, 2021, MDPI.



# Biography

Filip Turčinović was born in Rijeka in 1994. He completed his elementary education in Sveti Petar u Šumi (Vladimir Nazor Elementary School) and his secondary education in Pazin (Juraj Dobrila High School). In 2013, he enrolled at the Faculty of Electrical Engineering and Computing, University of Zagreb (FER). He successfully defended his undergraduate thesis in 2016 under the guidance of Professor Josip Knezović, and in 2018, he completed his master's thesis under the supervision of Professor Sonja Grgić. In November 2018, he began working as a researcher at the Department of Communication and Space Technologies at FER on the project "Development of an open smart network for energy-efficient public LED lighting," led by Professor Gordan Šišul. In October 2019, he embarked on a PhD program at FER under the mentorship of Professor Marko Bosiljevac. His research interests are closely linked to the project "Smart platform for millimeter-wave environmental and industrial sensing," on which he has been employed since October 2020. Together with Professor Dario Bojanjac, he founded the Institute for Science Outreach. He has authored a total of 16 scientific papers, four of which have been published in international journals. In recognition of his collaborative scientific and artistic efforts, he received the Rector's Award in 2018.