Strategos, 7(2), 2023, 41-58
UDK 32
UDK 355/359
Original scientific paper

STRATEGOS

Application of neural networks for the detection and classification of artillery targets

Adriana Stubičar, Mario Šipoš

Abstract

Neural networks have been in use since the 1950s and are increasingly prevalent in various domains of human activity, including military applications. Notably, GoogLeNet and convolutional neural networks, when appropriately trained, are instrumental in identifying and detecting individual objects within a complex set. In military scenarios, neural networks play a crucial role in the fire support process, especially when receiving target descriptions from forward observers. These networks are trained on image datasets to recognize specific features of individual elements or military objects, such as vehicles. As a result of this training, when presented with a new image, the network can accurately determine the type of vehicle, expediting the targeting process and enhancing the ability to provide a suitable response. This paper describes the application of neural networks for detecting and classifying artillery targets. It presents a specific problem and proposes a scientific solution, including explaining the methodology used and the results obtained.

Keywords

convolutional neural network, target detection, data augmentation

Introduction

Military target detection has been benefitting greatly from the growing popularity of convolutional neural networks (CNNs). CNNs are able to extract features and classify images with impressive accuracy, making them an ideal tool for this application. Specifically, CNNs are essential for synthetic aperture radar (SAR) target detection, a critical application in military and civilian domains. SAR images are known for their complexity, requiring sophisticated algorithms for target detection. Traditional target detection techniques that rely on the constant false alarm rate (CFAR) principle are insufficient in addressing the nuances of SAR imagery. Therefore, there is a pressing need for innovative approaches to solve SAR target detection problems, and CNNs are promising solutions to meet this demand (Huang *et al.*, 2021).

CNNs are efficient in military target detection due to several advantages they offer. Specifically, CNNs can acquire fine bounding boxes and the related class of objects, which is crucial in military operations. Additionally, CNNs can capture spatial, temporal, and thermal data and integrate them into a three-channel image, which can be combined as CNN highlight maps unsupervised. The Quick R-CNN technique inspires this technique and uses a cross-space move learning strategy to fine-tune the CNN model on multi-channel images. CNN architectures are commonly used in object recognition in the military, and transfer learning techniques can be applied to extract features learned from a large dataset and transfer them to another CNN architecture (Janakiramaiah *et al.*, 2023).

Military target detection can significantly benefit from the use of CNNs. These powerful tools can accurately extract and classify features from images. The advantages of CNNs in this context are numerous, such as their capacity to handle large amounts of data, high accuracy, and ability to learn and adapt to new data. CNNs have proven helpful in various military target detection scenarios, including SAR target detection, ship detection, and military object detection.

Neural networks

Neurons are the basic functional unit of the nervous system responsible for transmitting information between different body parts. Neurons have several components, including the cell body, dendrites, and axons. The dendrites receive input from other neurons or sensory organs, while the cell body integrates this information and sends the resulting output signal down the axon. Neurons communicate with each other at synapses, which are specialized sites of contact between cells where neurotransmitters are released. The basic properties of neurons and their signaling pathways are highly conserved across many species, and even simple systems like those found in the nervous system of invertebrates can give us valuable insights into the workings of more complex brains, including our own (Nicholls *et al.*, 2012).

Artificial neural networks are models used for machine learning that are inspired by the structure and function of the human brain. The network consists of interconnected nodes, or neurons, where each neuron takes in one or more inputs, applies a mathematical transformation, and produces an output. These outputs can be used as inputs for other neurons in the network, and the connections between neurons are weighted, meaning that some inputs have more influence on the output than others. During training, the model is given examples with inputs and corresponding desired outputs. Backpropagation is typically used to calculate the gradient of the loss function concerning the network's weights and then adjust the weights in the direction of the negative gradient. The loss function measures the difference between predicted and desired outputs. As the network is trained on more examples, it can learn to generalize to new inputs it has not seen before. The reason neural networks are so versatile is their ability to generalize, which enables them to perform tasks such as recognizing images, processing natural language, and understanding speech. Nonetheless, the complexity of deep learning models necessitates significant amounts of data and computational resources to be trained and used effectively (Goodfellow, Bengio and Courville, 2016).

Convolutional neural network (CNN) is a type of neural network that is especially well-suited for handling data with a grid-like structure, such as images and speech signals. At a high level, CNNs consist of layers of simple processing nodes that receive inputs, transform them using linear algebraic operations, and pass the outputs to subsequent layers. The main difference between CNNs and other types of neural networks is that they incorporate convolutional layers and pooling layers, which are designed to capture local image features and reduce dimensionality. In the convolutional layer, sets of weights, or kernels, slide or convolve over regions of the input image, transforming those regions into corresponding regions of an output feature map. The kernels capture local image features, such as edges, corners, and blobs, and each output feature map encodes a different subset of those features. The convolutional operation is typically followed by a ReLU activation function, which sets negative values to zero, producing a sparse output that is easier to compute and generalize while retaining non-linearities in the data.

In the pooling layer, the network reduces the dimensionality and increases the processing efficiency by performing local downsampling, typically by either max pooling or average pooling. This operation extracts the maximum or average response of each kernel output over a local region of the feature map, pooling similar features and reducing the overall number of parameters in the network. In addition to these layers, CNNs can include other types of layers, including dropout, batch normalization, and fully connected layers. Dropout layers can help prevent overfitting by randomly setting a fraction of the input units to zero during training, while batch normalization layers help normalize the input data to the subsequent layers. Finally, the fully connected layers transmit the high-level features learned by the convolutional and pooling layers to a final output layer for classification. CNNs have shown remarkable success in various applications, including image classification, object detection, facial recognition, and speech recognition. By automatically learning features from data, CNNs have reduced the need for manual feature engineering and have demonstrated strong generalization ability on unseen data (Patterson and Gibson, 2017).

Related work

Detecting and classifying military objects using CNNs has been a research subject for some time. CNNs effectively identify and categorize military objects, including vehicles, ships, and ground targets. There are several specific aspects of this research, such as:

Bilecik *et al.* (2022) proposed a classification approach based on a Mask Region-based Convolutional Neural Network (CNN) for detecting vehicles in UAV images. The authors discuss the growing use of UAV images in various fields, such as agriculture, security, the military, and highlight the importance of vehicle detection in military operations. The study shows that deep learning techniques, particularly CNN-based segmentation algorithms, can detect objects in images effectively. The authors demonstrate the effectiveness of Mask R-CNN, a CNN architecture-based approach, in detecting vehicles with high accuracy in images taken by UAVs. The paper provides an insightful discussion of the advantages of deep learning techniques, specifically Mask R-CNN, and their potential applications in other fields.

Kafedziski, Pecov and Tanevski (2018) proposed a solution for automated landmine detection using a Faster R-CNN with a Ground Penetrating Radar (GPR) system. The method uses deep learning and convolutional neural networks to discriminate between anti-tank (AT) mines and other objects. The system works well with custom datasets and can be extended to an arbitrary number of classes. The authors evaluate the performance of the method using Confusion matrices and ROC curves. The training and test datasets consist of both simulated and real measured data. The proposed method is considered a more efficient and safer alternative to traditional land mine detection methods.

Wu *et al.* (2018) discuss using the YOLO (You Only Look Once) network for object detection in remote sensing imagery. The YOLO network structure, which has 24 convolutional layers and two fully connected layers, is used for detecting objects in the images. The authors have pre-trained these convolutional layers on the ImageNet classification to enhance detection resolution. The paper also highlights the significance of data augmentation in improving the robustness of the neural network. For data augmentation, the authors have employed various image operations like rotating 45 degrees, scaling 15-25%, cutting, switching frequency bands, and vertical/horizontal flipping. The authors have utilized the Adam optimizer to optimize the model parameters. The Adam optimizer adjusts the learning rate of each parameter according to the first-moment estimation and the second-moment estimation of its gradient function of the loss function.

Deepthi, Kumar and Suresh (2021) proposed a custom Convolutional Neural Network (CNN) to detect and classify objects in satellite images. The need for automation of object detection lies in the fact that manual identification of objects in satellite images is very time-consuming, and it becomes difficult to manage large volumes of data. Using a custom CNN, this work addresses the challenge of automated detection and classification of objects such as trees, buildings, and cars in satellite images. The proposed network was tested on satellite images and showed promising results with high detection accuracy. The authors have discussed the performance characteristics of the custom CNN, highlighting that it provides superior results to other machine learning techniques.

Kim, Song and Kim (2017) provided a technical paper discussing the problem of infrared (IR) target recognition using deep convolutional neural networks (CNNs). The paper proposes two strategies to deal with the problem of limited IR image data and target variations, which hinder CNN training. The first strategy is to train CNN models using all possible examples, similar to existing models, representing the RGB domain. However, this approach is not feasible due to the limited number of IR images. Therefore, the second strategy preprocesses IR images in the feature space using edge information or aggregated channel feature (ACF) to reduce IR variation and improve recognition rates. The paper describes how edge-CNN can accurately perform IR target recognition using a gradient magnitude map, gradient orientation map, or ACF-CNN. The paper also covers SE-ATMOSPHERE, which generates atmospheric data, and SE-SCENARIO, which controls targets and backgrounds to help generate synthetic IR images for CNN training. The paper uses several figures showing experimental IR target recognition and detailed technical diagrams to support their proposals. The results of evaluating simulated infrared images using the thermal simulator (OKTAL-SE) have demonstrated the usefulness of IVO-CNN in military Automatic Target Recognition (ATR) applications.

Chen *et al.* (2019) presented an improved object detection and segmentation method based on Convolutional Neural Network (CNN) for military applications such as reconnaissance and strike. The proposed method adds Fisher discriminative criterion after the classification branch's Region of Interest (ROI) pooling layer. The proposed model has three branches: classification, regression, and mask segmentation. This paper describes the framework of object detection based on their proposed model, which includes data acquisition, model training, and detection stages. The authors use a Visible Image Dataset of multi-category armored targets in multiple scenarios acquired by an image acquisition device for model training. The effectiveness of the improved method is verified on two datasets, COCO and multi-category armored targets. The experimental results show that their proposed method can improve the accuracy of multi-category target detection and can accurately segment the target.

Tayara and Chong (2018) proposed a deep learning model for detecting and identifying objects in high-resolution aerial imagery. The model utilizes a top-down densely connected feature pyramid pathway to improve the detection and classification performance while reducing computation time. The authors conduct extensive experiments on a public dataset, comparing the performance of their model with various state-of-the-art models. The results showed that the proposed model achieves superior performance in terms of both accuracy and computation time. The authors suggest that their model has applications in urban planning, defense, and military applications.

Janakiramaiah *et al.* (2023) introduced a novel approach to enhance military object recognition using a relatively recent neural network architecture known as Capsule Network (CapsNet). Furthermore, they present a modified version of CapsNet, the multi-level CapsNet framework, aimed at improving the efficiency of military object recognition in scenarios where training datasets are limited. The effectiveness of the model under consideration is substantiated by an assessment using a dataset of images of military objects. This evaluation entails a comparative analysis against the predictions generated by support vector machines and Convolutional Neural Network (CNN) architecture. The dataset used in the experiments contained 3500 images of military and general objects collected online. The performance of the proposed techniques was evaluated using a ten-fold cross-validation method. The proposed deep transfer learning technique performs relatively better than other techniques reported in the literature.

GoogLeNet

GoogLeNet represents a prominent convolutional neural network (CNN) architecture developed by Google researchers in 2014. The primary aim of creating GoogLeNet was to strike a balance between computational efficiency and accuracy, ultimately aiming to improve image classification performance compared to the prevailing standards during that period. A distinctive feature of the GoogLeNet architecture is incorporating Inception modules. These modules are meticulously designed to capture features within the same network layer at various scales and resolutions. Each Inception module comprises a series of parallel convolutional layers with distinct filter sizes. A pooling layer subsequently follows these layers, and the outputs from each parallel layer are concatenated. This intricate setup enables the network to distinguish between fine-grained and coarse-grained features within a single layer, thus improving its image classification capabilities.

In addition to the Inception modules, GoogLeNet includes several other design innovations, contributing to its exceptional performance. It adopts global average pooling instead of fully connected layers at the termination of the network, which diminishes the number of parameters and mitigates overfitting issues. Moreover, GoogLeNet incorporates auxiliary classifiers at the intermediate layers of the network. These auxiliary classifiers stimulate the network to learn more discerning features and enhance its generalization capacity to novel images. In the broader context of deep learning, GoogLeNet marked a significant advancement and underscored the importance of creating neural network architectures that are accurate and computationally efficient. Its success has been influential, inspiring the development of subsequent architectures, including later iterations of the Inception architecture and other renowned networks such as ResNet and VGGNet (Szegedy *et al.*, 2015).

Training and validation

Data augmentation is a methodology used in machine learning to increase the size of a training dataset by generating modified versions of existing data. This technique is valuable when collecting high-quality data proves challenging, and its advantages extend to various fields of study. One notable benefit is the enhancement of model robustness and performance. Below, we elaborate on distinct data augmentation techniques:

1. Zooming: Adjusting the zoom level of an image allows the creation of new images with varying levels of detail, facilitating the ability of the model to identify objects at different scales and improving its scalability and adaptability to diverse scenarios.
2. Cropping: Cropping an image produces fresh images with different compositions. This is instrumental in helping the model understand objects in diverse contextual settings, thereby contributing to its contextual awareness.
3. Rotation: Rotating an image introduces new images with altered orientations. Consequently, the model gains an improved capacity to recognize objects from various angles, thus broadening its perspective.
4. Scaling: Altering the scale of an image generates images of varying sizes, enabling the model to recognize objects situated at different distances, strengthening the distance-related recognition capabilities of the model.
5. Horizontal flipping involves reversing the left and right sides of the image to create a flipped version. This technique helps improve symmetry and balance in images, correcting the orientation of an image captured or scanned in the wrong direction and generating new visual perspectives and variations.

6. Shearing: The process of shearing an image generates images of different shapes, assisting the model in recognizing objects with different aspect ratios and enhancing its adaptability to objects of varying proportions.
7. Channel shifts: Shifting the RGB or HSV channels within an image leads to new images with altered color distributions, equipping the model with the ability to recognize objects under diverse lighting conditions expanding its scope of adaptability.
8. Contrast adjustment: Adjusting the contrast of an image results in new images with varying levels of brightness and darkness, improving the ability of the model to identify objects under different lighting conditions and making it more versatile.
9. Noise: Introducing noise to an image produces images with different noise levels, equipping the model to recognize objects in noisy environments and enhancing its performance in less-than-ideal conditions.
10. Blurring: Blurring an image leads to creating images with varying sharpness levels, sharpening the proficiency of the model in recognizing objects within blurry images making it more adaptable.

These data augmentation techniques can be applied individually or in combination, effectively generating many new images from a limited dataset. By employing data augmentation, models can achieve enhanced performance and reduced overfitting risks, ultimately enhancing their overall effectiveness. In our research, we used zooming, cropping, rotation, scaling, and horizontal flipping to increase the number of photos for training CNN. The classification is based on the number of used images, ranging from 20 to 200. After data augmentation, the number of images is from 100 to 1000.

Figure 1 shows the relationship between the number of images used for CNN training and the accuracy of the model after training, represented in percentages. When using a CNN for training, the number of images used affects the accuracy of model after training. The percentage of accuracy improves as more images are used, with 66% accuracy achieved when 20 images are used for training. When 40 images are used, the accuracy improves to 70%. Similarly, the accuracy is 74% for 60 images, 76% for 80 images, and 80% for 100 images. Using a more significant number of images for training leads to further improvements in accuracy. The accuracy continues to improve, with 82%, 86%, 88%, and 90% accuracy obtained when using 120, 140, 160, and 180 images. Ultimately, when 200 images are used for training, the model achieves an accuracy of 92%.

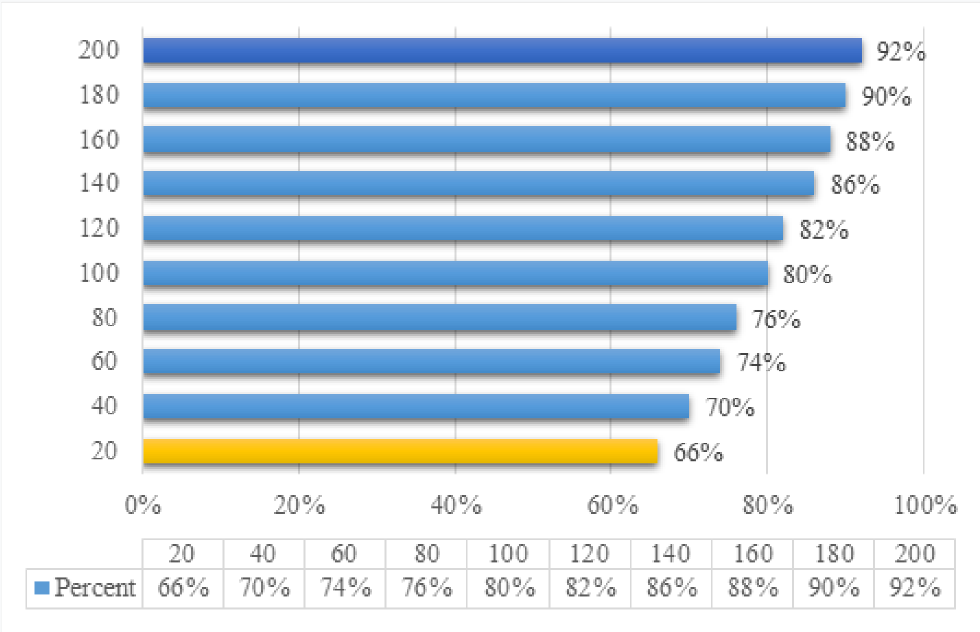

Figure 1. The relationship between the number of images used for CNN training and the accuracy of the model after training

Table 1 shows the results of a neural network model classifying two types of military vehicles, Abrams and Patria. The table includes ten rows, each representing a set of ten images. Each column represents the classification results.

The 'Y' sign indicates that the model classified the image correctly, while the 'N' sign means otherwise. For the Abrams class, images in rows 2, 3, and 5 were classified correctly in all dataset sizes. For the Patria class, images in rows 7 and 8 were classified correctly in all dataset sizes. In contrast, images in rows 6, 9, and 10 for Patria and row 1 for Abrams were misclassified for most dataset sizes. The table gives valuable information about the accuracy rates of the model as the number of trained images increases.

Table 1. Results of a neural network model classifying two types of military vehicles.

		Dataset									
Number	Image	20	40	60	80	100	120	140	160	180	200
1	Abrams	N	N	N	N	N	Y	Y	Y	Y	Y
2	Abrams	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3	Abrams	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4	Abrams	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
5	Abrams	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
6	Patria	N	N	N	N	N	Y	Y	Y	Y	Y
7	Patria	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
8	Patria	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
9	Patria	N	N	N	N	N	N	N	N	N	Y
10	Patria	N	N	N	N	N	N	N	N	N	Y

Table 2 shows the results obtained by learning the GoogLeNet neural network. The table provides performance evaluation results based on the number of tested images, 40, 120, and 200, and the specific vehicle types, Patria and Abrams. A 'Patria' or 'Abrams' label indicates the actual vehicle class of the image, and the table indicates the degree of accuracy in recognizing the class of the image by the neural network. The results are provided in percentages, where higher percentages represent higher accuracy. The table thus provides valuable information about the performance of the GoogLeNet model for vehicle detection tasks.

Table 2. Results obtained by learning the GoogLeNet network.

		Results					
		40		120		200	
Number	Image	Patria	Abrams	Patria	Abrams	Patria	Abrams
1	Abrams	3.45	96.55	2.60	97.40	1.66	98.34
2	Abrams	44.80	75.20	22.45	77.56	16.15	83.85
3	Abrams	16.08	83.92	14.37	85.63	11.17	88.83
4	Abrams	38.74	61.26	31.18	68.82	28.41	71.59
5	Abrams	1.86	98.14	1.33	98.67	1.44	98.56
6	Patria	99.79	0.21	99.86	0.14	99.88	0.13
7	Patria	99.89	0.11	99.79	0.21	99.79	0.21
8	Patria	99.76	0.24	99.66	0.34	99.68	0.32
9	Patria	98.46	1.54	98.13	1.87	98.83	1.18
10	Patria	82.73	17.27	85.27	14.73	87.21	12.79

The significant oscillation in the correct recognition of armored vehicles from one image set to another or from one row of the table to another in Table 2 can be attributed to several factors, including the number of images used for training the neural network model. Table 2 provides performance evaluation results based on the number of training images, 40, 120, and 200, which can affect the accuracy of the model. The accuracy of the model may vary depending on the number of tested images due to variations in lighting, image quality, and object appearance that may exist among the images. These variations can cause inconsistencies in the performance of the model for object detection and recognition tasks.

A higher percentage of accuracy indicates a better performance of the model, but the acceptable level of accuracy would vary depending on the specific use case. For example, in a military context, a percentage of accuracy of 61.26 (Table 2, Number 4), which represents the accuracy of the model in recognizing Patria vehicles when tested with a sample of 40 images, may not be considered acceptable if critical decisions are being based on the output of the model. In contrast, a lower acceptable level may be sufficient for satisfactory performance in other contexts, such as in a basic object recognition.

Moreover, the acceptance level can also be influenced by various other factors, including the importance of the object being detected and classified, the criticality of the model's output, and the consequences of detection failures or false positives. Therefore, the acceptance level of accuracy for each dataset would be subject to the specific application, requirements, and context of the model.

Using transfer learning, GoogLeNet achieves higher accuracy with fewer training images than a neural network built from scratch. This technique involves utilizing a pre-trained model as a starting point for a new model. The pre-trained model has already learned to extract robust and informative features from natural images, and this knowledge can be transferred to the new model. Most pre-trained neural networks are trained on a subset of the ImageNet database, which is used in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). These neural networks have been trained on over a million images and can classify images into 1000 object categories. Using a pre-trained neural network with transfer learning is typically faster and easier than training a neural network from scratch.

Conclusion

Neural networks have found significant applications in the military, a fact recognized by several nations, including the United States. Furthermore, their relevance extends to artillery for target acquisition. A well-trained neural network can detect various military objects, such as vehicles, allowing for the identification and determination of their specific technical characteristics.

The use of neural networks should increasingly become visible within the Croatian military, especially in artillery, given their practical and vital applications. The effective implementation of neural networks accelerates the target detection and classification processes, consequently improving accuracy in hitting targets. When using this technology, unmanned aerial vehicles (UAVs) can be employed, or cameras can be strategically positioned, with oversight entrusted to the commanding officer.

Our research has shown that using data augmentation techniques such as zooming, cropping, rotation, scaling, and horizontal flipping can significantly improve the accuracy of convolutional neural networks (CNNs) for image classification. We observed that the accuracy of the trained model improves as the number of images used for CNN training increases. Ultimately, the model achieved an accuracy of 92% when 200 (1000) images were used for training. Furthermore, we observed that transfer learning techniques can boost the performance of CNNs by extracting features learned from a large dataset and transferring them to another CNN architecture. This approach can be particularly effective when dealing with limited training data.

Comparing GoogLeNet and trained CNN from scratch, we found that while GoogLeNet performed better on a smaller dataset, our trained CNN achieved comparable performance on a larger dataset using data augmentation techniques. This suggests that the choice of CNN architecture should be based on the size and quality of the available training data.

In the context of military target detection, CNNs offer numerous advantages, such as their ability to handle large amounts of data, high accuracy, and ability to learn and adapt to new data. CNNs have proven effective in various military target detection scenarios, including SAR target detection, ship detection, and military object detection.

References

Bilecik, Yayla, R., Albayrak, E., Bilecik Şeyh Edebali University, Turkey, Yüzgeç, U., and Bilecik Şeyh Edebali University, Turkey (2022). Vehicle Detection from Unmanned Aerial Images with Deep Mask R-CNN, *Computer Science Journal of Moldova*, 30(2 (89)), pp. 148–169. Available at: https://doi.org/10.56415/csjm.v30.09.

Chen, D., Yang, C., Liu, Z., Zhang, X. and Shi, S. (2019). Object Detection and Segmentation Method for Multi-category Armored Targets Based on CNN, in Y. Wang, Q. Huang, and Y. Peng (eds) *Image and Graphics Technologies and Applications*. Singapore: Springer (Communications in Computer and Information Science), pp. 61–69. Available at: https://doi.org/10.1007/978-981-13-9917-6_7.

Deepthi, S., Kumar, S. and Suresh, L. (2021). Detection and Classification of Objects in Satellite Images using Custom CNN, *International Journal of Engineering Research*, 10(06).

Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*.

Huang, Q., Zhu, W., Li, Y., Zhu, B., Gao, T. and Wang, P. (2021). Survey of Target Detection Algorithms in SAR Images, in *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Chongqing, China: IEEE, pp. 1756–1765. Available at: https://doi.org/10.1109/IAEAC50856.2021.9390728.

Janakiramaiah, B., Kalyani, G., Karuna, A., Prasad, L.V.N. and Krishna, M. (2023). Military object detection in defense using multi-level capsule networks, *Soft Computing*, 27(2), pp. 1045–1059. Available at: https://doi.org/10.1007/s00500-021-05912-0.

Kafedziski, V., Pecov, S. and Tanevski, D. (2018). Detection and Classification of Land Mines from Ground Penetrating Radar Data Using Faster R-CNN, in *2018 26th Telecommunications Forum (TELFOR)*. *2018 26th Telecommunications Forum (TELFOR)*, Belgrade: IEEE, pp. 1–4. Available at: https://doi.org/10.1109/TELFOR.2018.8612117.

Kim, S., Song, W.-J. and Kim, S.-H. (2017). Infrared Variation Optimized Deep Convolutional Neural Network for Robust Automatic Ground Target Recognition, in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2017* Honolulu, HI, USA: IEEE, pp. 195–202. Available at: https://doi.org/10.1109/CVPRW.2017.30.

Nicholls, J.G., Martin, A.R., Fuchs, P.A., Brown, D.A., Diamond, M.E. and Weisblat, D.A. (2012). *From Neuron to Brain*. Fifth Edition.

Patterson, J. and Gibson, A. (2017). *Deep Learning A Practitioner's Approach*.

Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. (2015). Going deeper with convolutions, in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, pp. 1–9. Available at: https://doi.org/10.1109/CVPR.2015.7298594.

Tayara, H. and Chong, K. (2018). Object Detection in Very High-Resolution Aerial Images Using One-Stage Densely Connected Feature Pyramid Network, *Sensors*, 18(10), p. 3341. Available at: https://doi.org/10.3390/s18103341.

Wu, Z., Chen, X., Gao, Y. and Li, Y. (2018). RAPID TARGET DETECTION IN HIGH RESOLUTION REMOTE SENSING IMAGES USING YOLO MODEL, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII–3, pp. 1915–1920. Available at: https://doi.org/10.5194/isprs-archives-XLII-3-1915-2018.

About the authors

First Lieutenant ADRIANA STUBIČAR (e-mail: adrianastubicar43@gmail.com) works as analyst at Ministry of Defenc. She is also studying at the postgraduate specialist study, Faculty of Mechanical Engineering and Shipbuilding in Zagreb, Industrial engineering and management.

Captain MARIO ŠIPOŠ, PhD (e-mail: mario.sipos86@gmail.com) works as an officer for science and development in the Technical Research and Development Department at the Center for Defense and Strategic Studies, Croatian Defense Academy.

Primjena neuronskih mreža za otkrivanje i klasifikaciju topničkih ciljeva

Sažetak

Neuronske mreže u uporabi su od pedesetih godina prošlog stoljeća i sve su zastupljenije u različitim područjima ljudske aktivnosti, uključujući vojne primjene. Posebno, GoogleNet i konvolucijske mreže, kada su odgovarajuće utrenirane, ključne su u prepoznavanju i otkrivanju pojedinih objekata unutar složenog skupa. U vojnim scenarijima neuronske mreže imaju ključnu ulogu u postupku pružanja potpore vatrom, posebno kada primaju opise ciljeva od prednjih topničkih motritelja. Ove mreže trenirane su na slikovnim skupovima podataka kako bi prepoznale specifičnosti pojedinih elemenata ili vojnih objekata, kao što su vozila. Kao rezultat treniranja, kada se prikaže nova slika, mreža može točno odrediti tip vozila, ubrzati postupak ciljanja i poboljšati sposobnost pružanja prikladnog odgovora. U radu se opisuje primjena neuronskih mreža za otkrivanje i klasifikaciju topničkih ciljeva. Rad predstavlja poseban problem i predlaže rješenje primjenom znanstvenog pristupa, uključujući objašnjenje korištene metodologije i dobivene rezultate.

Ključne riječi

konvolucijska neuronska mreža, otkrivanje ciljeva, povećanje podataka