2025

Monograph

Zoriana Rybchak, Oleh Basystiuk

INTELLIGENT METHODS FOR BUILDING AN ONTOLOGICAL KNOWLEDGE BASE FOR CLASSIFYING THERMAL PORTRAITS OF TARGETS

ScientificWorld
Germany

Zoriana Rybchak, Oleh Basystiuk

INTELLIGENT METHODS FOR BUILDING AN ONTOLOGICAL KNOWLEDGE BASE FOR CLASSIFYING THERMAL PORTRAITS OF TARGETS

In internationalen wissenschaftlich-geometrischen Datenbanken enthalten
Included in International scientometric databases

MONOGRAPH

ScientificWorld-NetAkhatAV
Karlsruhe 2025

Authors:
Zoriana Rybchak, Oleh Basystiuk

Intelligent methods for building an ontological knowledge base for classifying thermal portraits of targets: Monograph / Z. Rybchak, O. Basystiuk. Karlsruhe, 2025. 104 p.
Monographic series «Innovative Science, Education, Manufacturing and Transport». Book 19. 2025.

ISBN 978-3-98924-119-0
DOI: 10.30890/978-3-98924-119-0.2025

Published by:
ScientificWorld-NetAkhatAV
Lußstr. 13
76227 Karlsruhe, Germany
e-mail: *editor@promonograph.org*
site: https://de.promonograph.org

ABOUT THE AUTHORS

1. *Zoriana Rybchak,* PhD, Associate Professor, LPNU
ORCID: 0000-0002-5986-4618

2. *Oleh Basystiuk,* PhD, LPNU
ORCID: 0000-0003-0064-6584

Content

 ISBN 978-3-98924-119-0

Introduction

Ontologies are a modern direction in the field of artificial intelligence. In the general sense, ontology can be characterized as a special type of knowledge base or as a “specification of conceptualization” of a particular subject area. This implies that within a specific location, based on the classification of basic terms, key concepts are highlighted and the relationships between them are determined. In the future, the ontology can be represented in the form of a graphical model or described using formal languages, which is the process of creating an ontology specification. Ontological knowledge bases are a powerful tool for organizing and processing large amounts of information. They allow you to formalize knowledge in the form of a clear structure, where each element of the base has a clearly defined meaning, properties, and relationships with other components. They are handy for working with data that has a complex hierarchical or relational structure, as is often the case when analyzing thermal portraits of targets.

The development of an ontological knowledge base of thermal portraits of targets is a crucial and advanced step in the development of modern technologies, with a broad range of applications across various fields of human activity, making it a highly relevant task.

In the context of modern military conflicts, the effective detection and identification of targets are key aspects for the successful performance of combat missions. In particular, in conditions of limited visibility, such as nighttime, fog, or smoke, traditional observation methods become ineffective. Thermal imaging technologies enable the detection of thermal signatures from objects, significantly improving the accuracy of determining the position and characteristics of targets. The development of an ontological knowledge base of thermal portraits of military equipment enables the creation of a generalized and structured representation of knowledge about the thermal properties of various types of equipment, contributing to faster and more accurate target identification.

Due to the intensive development of WMD, there is a need to create new systems

 ISBN 978-3-98924-119-0

for their identification, since classical methods no longer provide the necessary level of reliability and efficiency in solving this problem. The development of a new approach to WMD identification is a significant and promising task in enhancing Ukraine's defense capabilities. For such identification, it is necessary to develop an ontological knowledge base of thermal portraits of targets. The primary advantage of infrared thermal imaging cameras is their ability to visualize the surrounding world and specific objects, regardless of illumination levels (even in complete darkness) and weather conditions.

Thanks to the use of machine learning technologies based on the developed ontology, it is possible to develop a software module for identifying targets using panoramic thermal imaging cameras. This information system will be a competitive analogue to systems created on the global market and will also have specific advantages over existing systems, such as the accuracy of automatic location determination (through the use of the 3D triangulation method of infrared images) and higher performance (through optimal adjustment of the thermal map update period in the ontological knowledge base).

It should be added that, when analyzing the promising development of a new generation of tanks in the world's leading countries, the following main design features can be identified:

- The design and configuration of tanks provide for the placement of modern reconnaissance and strike capabilities.
- Tank improvements are based on modular design solutions, automation of fire control, protection, and mobility processes, increased survivability and autonomy, improved ergonomics, and the introduction of artificial intelligence elements into tank designs.
- Equipped with powerful weapon systems, night vision devices, laser rangefinders, angle measurement systems, laser beam control channels, integrated sights with five optical channels, and different operating wavelengths, and thermal imaging sighting systems.

Today, thermal imaging sighting systems are widely used in the fire control

ISBN 978-3-98924-119-0

systems of tanks, infantry fighting vehicles, and armored personnel carriers in the armies of the world's leading countries.

In the modern world, new technologies are developing rapidly, and their implementation in various fields of activity, particularly in the military, is becoming increasingly important. In particular, this concerns the use of modern services and tools for recognizing enemy weapons and military equipment. Such technologies can significantly improve the efficiency of reconnaissance, monitoring, and control of combat operations. One of the key areas is the use of advanced sensor technologies, including thermal imagers, infrared cameras, drones equipped with high-precision sensors, and crewless aerial vehicles.

They enable surveillance of the territory, detection of moving objects, and identification of thermal signatures characteristic of military equipment. The integration of such technologies into military strategies enables the acquisition of accurate information about the enemy’s location and the timely detection of potential threats. In addition, the latest services based on artificial intelligence and machine learning provide automatic recognition of objects in thermal images and video, allowing you to quickly identify types of equipment and determine their characteristics, such as engine temperature, dimensions, and other parameters. This significantly reduces the time required for data analysis, allowing you to make decisions in real-time quickly.

By integrating such tools into intelligence and control systems, it is possible to significantly enhance the accuracy and speed of response to battlefield changes, thereby increasing the chances of success in combat conditions. Military systems using these technologies can effectively monitor large areas, as well as interact with other combat units, maintaining synchronization of actions and providing operational support in high-stress conditions. Thus, the latest technologies are becoming indispensable tools in the military sphere, contributing not only to improving the effectiveness of surveillance but also to ensuring higher accuracy in recognizing and identifying enemy equipment. They open up new opportunities for analysis and decision-making, which is a crucial factor in achieving a strategic advantage.

Methods and means of creating ontologies can be used to detect, analyze, and recognize thermal portraits of targets, particularly for identifying and monitoring enemy weapons and military equipment. In this context, a key feature is the high speed of processing thermal images, which enables a timely response to threats.

Information systems methods aimed at developing and creating systems capable of automatically collecting, processing, and classifying data from various sources (thermal imagers, radars, optical cameras). This includes creating architectures for systems that integrate machine learning and deep learning technologies for automated analysis of thermal images.

Ontology engineering involves creating ontologies for structuring knowledge in the form of graphical models or formal languages. This enables you to organize and classify data, enhancing the efficiency of analyzing thermal portraits of targets and providing a high level of integration and automation of processes.

Data analysis uses statistical and machine learning methods to process large volumes of thermal images. This includes the use of filtering, classification, and data integration methods to achieve high accuracy of object identification.

Deep learning methods utilize neural networks for automatic learning on large datasets, allowing systems to adapt to new types of objects and changes in observation conditions. Particular attention is paid to the use of convolutional neural networks, which provide high accuracy of thermal image classification.

Artificial intelligence technologies integrate and automate the analysis of thermal portraits. This includes the use of algorithms for learning, optimization, and prediction, which allows for achieving high accuracy of target detection under conditions of limited visibility.

The project proposes a new approach to developing an ontological knowledge base of thermal portraits, which enables not only the collection and systematization of data on various types of equipment but also the provision of their in-depth understanding and interpretation through the integration of advanced artificial intelligence methods. Such a system allows for the automatic identification of military objects by their thermal characteristics, determines their type, size, and location, and

 ISBN 978-3-98924-119-0

predicts their behavior based on the analysis of past patterns. In addition, the use of ontology enables the effective processing of large volumes of data in real-time, ensuring a high speed of response to changing situations on the battlefield.

CHAPTER 1
APPROACHES TO THE CLASSIFICATION OF THERMAL PORTRAITS OF MILITARY TARGETS

The development of thermal portraits of targets is an important aspect in security systems, military technologies, energy, and medical diagnostic systems. The use of an ontological knowledge base allows for the structuring and automation of thermal data analysis, ensuring faster identification of objects based on their thermal characteristics and increased analysis accuracy through the integration of a large amount of knowledge.

With the development of technology, there is a growing demand for intelligent surveillance systems that provide a high level of automation and accuracy in various fields. One important area is the development of thermal portraits of targets, which have a wide range of applications in security and military technologies.

Thermal portraits are unique to each object, as temperature and thermal characteristics depend on various factors, such as the physical properties, environment, and activity of the object. The use of thermal imagers to create these portraits allows objects to be identified, observed, and analyzed even in conditions of limited visibility, such as darkness, smoke, or fog.

Thermal portraits of targets are based on unique thermal characteristics that can be used to identify objects. Each object—whether a person, vehicle, building, or other structure—has its own specific temperature properties. These properties depend on a number of factors, such as:

- Physical properties of the object: different materials from which objects are made have different abilities to retain heat, which creates specific thermal reflections.
- Environment: Ambient temperature and weather conditions such as wind, humidity, or solar heat can affect the temperature of an object and, accordingly, its thermal characteristics.
- Object activity: The activity of an object, such as movement or processes related to its operation, can change its thermal profile, for example, heating

 ISBN 978-3-98924-119-0

up from the operation of engines or human activity.

These factors together create a unique thermal profile, or thermal portrait, which can be captured and analyzed using thermal imaging cameras.

A military thermal imager is a specialized optical device designed to detect, monitor, and analyze objects using their thermal radiation. In modern combat conditions, where speed and accuracy of decision-making are important, such devices have become indispensable. They enable military units to conduct reconnaissance, detect the enemy, and observe the terrain in conditions of limited visibility, particularly at night.

The main purpose of a military thermal imager is to detect thermal signals emitted by people, equipment, or other objects. The device is capable of displaying temperature contrasts on the screen, which makes it possible to quickly respond to changes in the situation. Modern thermal imager models can be integrated with other surveillance systems, which significantly increases their effectiveness.

The principle of operation of a military thermal imager is based on the detection of infrared radiation emitted by objects depending on their temperature. All living organisms and equipment emit heat, which makes them visible to a thermal imager even in difficult conditions, such as thick smoke or darkness.

The thermal imager's operation involves several stages:

- Infrared radiation collection: the thermal imager captures the thermal signals emitted by objects.
- Signal processing: the received signals are processed by a processor, which converts them into a visual image.
- Visualization: a color image is formed on the screen, where warm objects are displayed in bright colors and cold objects in darker colors.

This process allows the military to quickly detect the enemy, even if they are at a great distance or in well-protected positions.

When choosing a military thermal imager, it is extremely important to pay attention to its technical characteristics, as they determine the effectiveness of the device in real combat conditions. One of the key parameters is resolution: the higher it is, the clearer and more detailed the image will be, and thus the operator's ability to recognize small targets and distant objects will increase. Equally important is the sensitivity of the thermal imager — an indicator that determines its ability to detect the slightest temperature fluctuations. In difficult situations, where the enemy may be camouflaged or partially obscured by obstacles, it is high sensitivity that makes it possible to detect the target.

Detection range is another key parameter: it shows the maximum distance at which the thermal imager can see an object, which is especially important for observation posts, reconnaissance, and patrolling the territory. The viewing angle is no less important. A wide field of view allows you to monitor a larger area, reducing the risk of missing danger from the flanks or rear. Military equipment must operate in rain, fog, dust, and temperature fluctuations, so protection from moisture and dust is critical. A sturdy housing with sealed joints ensures reliable operation of the device even in the worst weather conditions.

Thanks to their technical parameters, military thermal imagers have a number of advantages that make them indispensable in modern combat operations. They are not dependent on the level of illumination and are capable of operating in complete darkness, giving the military a significant advantage at night. The accuracy of thermal imagers makes it easy to detect the enemy due to their ability to detect the slightest temperature changes, which is especially important when observing enemy movements or searching for camouflaged objects. In addition, thermal imaging devices demonstrate high efficiency in difficult climatic and tactical conditions: smoke, heavy rain, snow, or fog have virtually no effect on their performance, whereas conventional optics become almost useless in such situations.

The versatility of thermal imagers allows them to be used in a wide range of military tasks, from surveillance and patrolling to supporting assault operations. Modern models have intuitive interfaces, allowing military personnel to quickly master their use in the field without wasting time on complex settings. As a result, thermal imagers have become important tools that significantly increase the response speed, decision-making accuracy, and overall effectiveness of military units.

 ISBN 978-3-98924-119-0

The scope of application of thermal imagers in military affairs is extremely wide. In reconnaissance operations, they allow the enemy to be detected at considerable distances, which increases situational awareness and helps to plan further actions taking into account the real situation. In special operations, where silence, speed, and minimal visibility are important, thermal imaging cameras are indispensable, allowing soldiers to avoid ambushes and quickly find key targets. For monitoring territory, thermal imaging cameras provide constant surveillance of strategically important objects, perimeters, and equipment, and also help prevent intrusions.

Their use in the training of military personnel occupies a special place. Thermal imagers allow you to simulate various tactical situations, practice observation, detection, and assessment skills, which increases the level of personnel training. Thus, military thermal imagers combine high technological capabilities, reliability, and versatility, making them key tools in modern defense infrastructure.

These areas confirm the importance of thermal imaging cameras for modern military operations. Today, the development of thermal imaging technology plays an important role in various fields, particularly in the military industry, where it is used to solve specific tasks related to monitoring and control. At the same time, there is a lack of systematic information , which limits the possibilities for a comprehensive and balanced solution to these tasks. In addition, there are different approaches to solving problems in the thermal imaging technology market, where experts are trying to find the optimal balance between technical capabilities and financial aspects.

Thermal imaging is based on the analysis of thermal radiation from bodies, which allows objects to be detected by their thermal field. Although this radiation is invisible to the eye, it can be detected using special thermal imagers. The spectrum of radiation emitted by objects varies, and IR radiation, due to its longer wavelength, has the ability to penetrate the atmosphere, but is also attenuated by various atmospheric obstacles such as water, carbon dioxide, ozone, and other gases.

There are two main classes of thermal imagers: cooled and uncooled. Cooled thermal imagers operate in the mid-wave range (3-5 μm) and require cryogenic cooling to ensure adequate sensitivity (Fig. 1.). They are used when high sensitivity to temperature differences and high resolution are required. However, they are expensive , require large amounts of energy, and have mobility limitations.

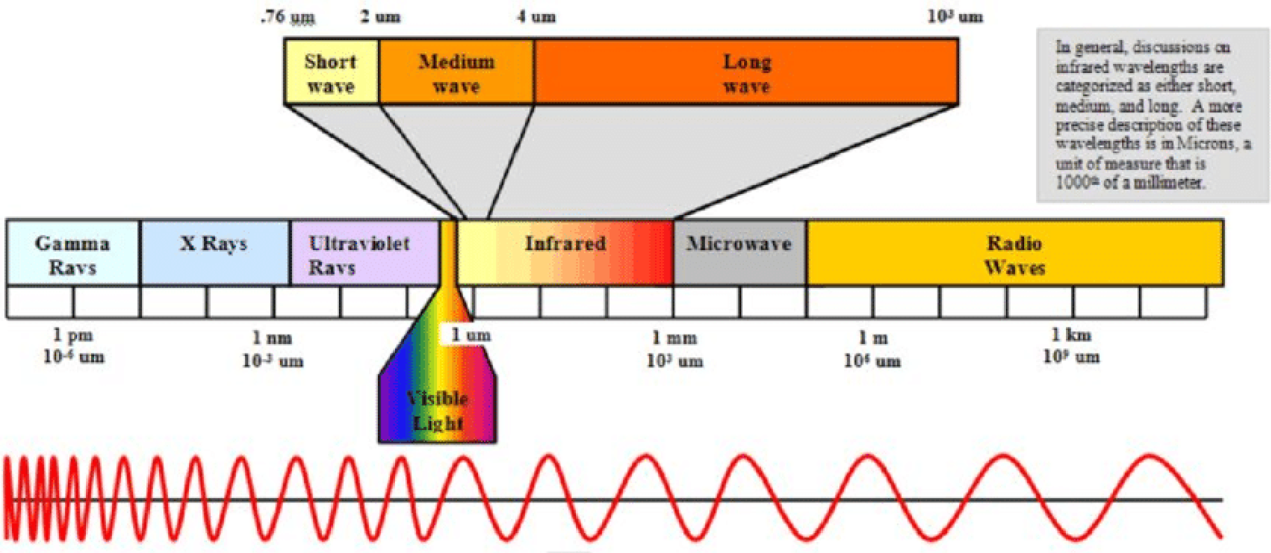

Fig. 1 - Spectra of visible and IR electromagnetic radiation

Uncooled thermal imagers operating in the long-wave range (7-14 μm) are less sensitive to temperature changes, but they are highly resistant to interference and natural atmospheric conditions such as smoke or fog. They are smaller, less expensive, and provide a faster transition to operating mode. However, to ensure the necessary efficiency, they require more powerful lenses, which increases the demands on the optics.

Thus, the choice between a cooled and uncooled thermal imager depends on the specific tasks. For many applications, such as medium-range surveillance, uncooled thermal imagers may be a more cost-effective option, as they provide sufficient performance for less money. However, for specialized tasks, particularly military applications such as aviation or long-range surveillance, cooled thermal imagers remain essential.

In Fig. 2, which shows LWIR (uncooled thermal imager) and MWIR (cooled thermal imager), you can see the effect of interference caused by engine exhaust and complex natural atmospheric conditions, which make it difficult for cooled thermal imagers to recognize objects. This effect is explained by the increased spectral sensitivity of cooled thermal imagers in the mid-wave range, which makes them more susceptible to the absorption of infrared radiation from gases such as carbon monoxide and nitrogen. On the other hand, uncooled thermal imagers, operating in the long-wave range, are less sensitive to such interference, since the absorption of these gases has a much smaller effect on transparency in this range. This makes uncooled thermal imagers more effective in poor atmospheric conditions, such as fog, smoke, or engine exhaust.

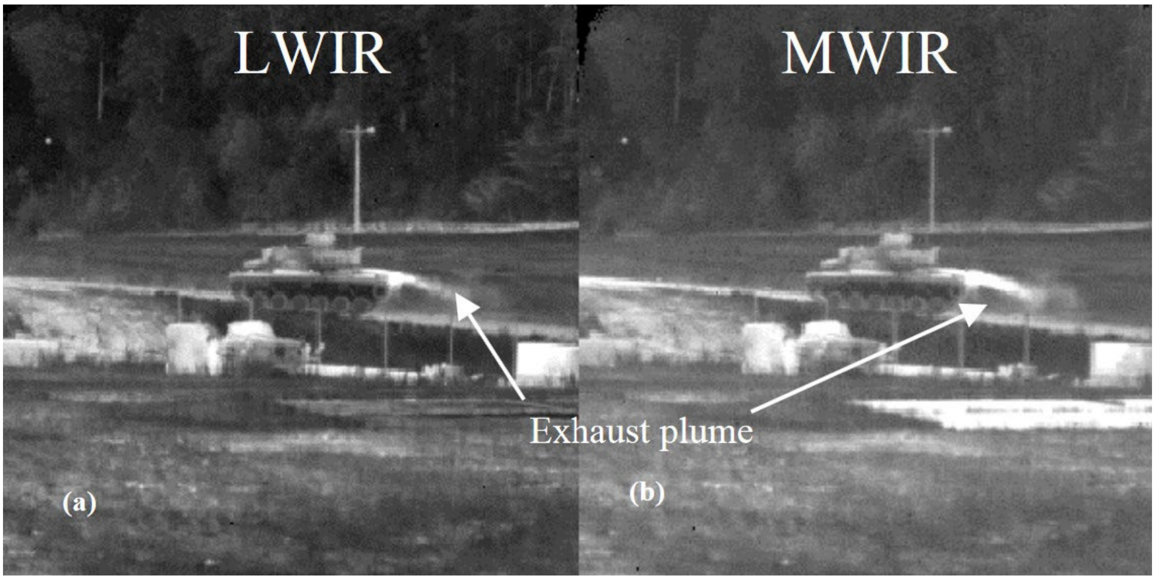

Fig. 2 - LWIR (uncooled thermal imager) and MWIR (cooled thermal imager) images

Cooled and uncooled thermal imagers differ significantly in design and component composition, and each has its own advantages and disadvantages, which determine their areas of application and economic feasibility in different fields. Cooled thermal imagers allow for greater distances for detecting and recognizing objects, but they are more sensitive to various interferences and difficult natural conditions. In addition, their high cost and need for regular maintenance make them less cost-effective for some applications.

In contrast, uncooled thermal imagers provide a viewing range of up to 10 km and are more resistant to external factors such as smoke, fog, or engine exhaust. They are significantly less expensive, making them more cost-effective for many applications.

The identification and classification of thermal portraits of targets, which can vary

 ISBN 978-3-98924-119-0

significantly in shape, intensity of thermal radiation and nature of spatial temperature distribution, is an extremely complex task. An additional problem is the variability of the quality of input data, in particular thermal or infrared images, which often contain noise, artifacts or limited resolution. This requires a sufficient amount of representative data and the use of intelligent processing methods capable of ensuring the correct semantic description of targets and the integration of knowledge within the ontological model.

In the context of thermal image processing, several basic types of computer vision tasks are traditionally distinguished:

- Classification;
- Localization;
- Object detection (one-stage and two-stage);
- Segmentation (semantic, instance, panoptic);
- Object tracking in dynamics;
- Classification of thermal images.

Classification involves determining the probability of a thermal portrait belonging to a certain category (for example: “person”, “heated vehicle”, “technical object”, “background infrastructure”). This method is the basic stage of the formation of ontological classes, as it allows identifying the key properties of the object necessary for further formalization in the ontology.

The localization task is aimed at determining a specific area of the image where the thermal object is located. This provides binding of semantic descriptions to spatial coordinates and allows to form ontological relations of the type “object is in the region”, “object crosses the segment boundary”, etc.

Unlike simple localization, detection simultaneously performs classification and determination of the location of the object, which makes it a universal tool for automated obtaining of structured facts that can be integrated into the knowledge base.

There are two main approaches:

1. Two-stage detection (R-CNN and derivatives) - involves preliminary formation of regions of interest, which provides high accuracy, important

for the formation of reliable elements of the ontological model.

2. One-stage detection (YOLO, SSD) - works faster, which allows it to be used in modes with limited computing resources or during streaming processing of thermal data, for example, in monitoring systems.

Segmentation allows not only to identify an object, but also to determine its shape and precise boundaries. This is critically important for building ontological descriptors that describe the structural and morphological characteristics of targets.

Main types of segmentation:

1. Instance - highlights the boundaries of each individual thermal object; allows you to create unique individual ontological representations.
2. Semantic - classifies each pixel, but without separating objects of the same class; useful for forming general ontological categories.
3. Panoptic - combines the two previous methods, providing the most complete semantic and spatial information to supplement the knowledge base.

Tracking is an extension of the detection task, as it allows you to describe the dynamic properties of targets. The obtained data can be used to create ontological relations such as “moves with a trajectory”, “changes thermal intensity”, “approaches/removes from the region of interest”. Methods based on the Kalman filter or neural network trackers are usually used.

Thus, for most tasks at "over-the-horizon" ranges (up to 7,200 m), for example, for sights and observation systems of linear combat vehicles (tanks, APCs, IFVs), infantry units, or anti-tank missile systems, uncooled thermal imagers are quite sufficient. They are also suitable for use in special operations forces and close reconnaissance. However, for tasks such as long-range reconnaissance, aviation, target designation, coordination of subordinate units, as well as for use in naval and riverine environments, where greater distance and accuracy of observation are required, it is more appropriate to use cooled thermal imagers.

The use of thermal imagers for the creation and analysis of thermal images plays a crucial role in a wide range of operational and analytical tasks. Thermal imaging

ISBN 978-3-98924-119-0

technology enables reliable identification of objects even in conditions of significantly reduced visibility, such as at night or in the presence of smoke, fog, or other atmospheric disturbances. This capability is particularly important in military environments, where early detection of personnel, equipment, or concealed positions can be decisive. Unlike conventional optical devices, thermal imagers do not depend on ambient light, which allows them to reveal targets that would otherwise remain hidden.

Another important application of thermal imaging is continuous surveillance. Thermal cameras make it possible to monitor extensive territories or specific objects without compromising operational stealth. They are therefore commonly used to secure perimeters of military bases, industrial zones, or critical infrastructure facilities such as oil and gas terminals. By capturing heat signatures, thermal sensors provide persistent situational awareness regardless of lighting conditions or obstacles that might interfere with optical systems.

Thermal imaging also supports advanced analysis of object behavior. Because thermal characteristics change depending on activity, these images provide valuable insights into both the state and dynamics of objects. For instance, an increase in the temperature of machinery can indicate its operation or overload, while variations in human thermal signatures may reflect movement, stress, or interaction with the environment. Thus, thermal portraits offer not only a snapshot of the object but also a means to interpret its ongoing behavior.

The advantages of using thermal portraits stem largely from their reliability and versatility. One of the primary benefits is the high accuracy they deliver in conditions where traditional observation tools are ineffective. This is vital in scenarios that demand immediate response and precise identification, such as military engagements, search-and-rescue operations, or emergency management.

Thermal imaging also supports a high degree of automation. When integrated with artificial intelligence and machine learning techniques, thermal surveillance systems can autonomously identify objects, classify targets, and detect anomalies by comparing new observations with existing knowledge bases. This facilitates real-time decision-

 ISBN 978-3-98924-119-0

making and reduces the need for constant human supervision.

Finally, thermal portraits are widely applicable beyond military and security contexts. In the energy sector, thermal imaging is used to detect heat loss or identify faulty components in power systems. In medicine, it enables non-invasive detection of temperature anomalies that may indicate inflammation or vascular disorders. Thus, the use of thermal imagers provides a universal and powerful tool for observation, analysis, and diagnostics across diverse domains.

The development of thermal portraits of targets using thermal imagers is an important area for improving the effectiveness of intelligent surveillance systems. They enable accurate identification, observation, and analysis of objects based on their thermal characteristics, even in conditions of limited visibility. Technologies that use thermal portraits are driving innovation in many areas, including security, defense, energy, and medicine.

The classification of thermal portraits of weapons and military equipment targets is an important area of military technology, as it allows for the effective identification and differentiation of various objects based on their thermal characteristics. The main innovative approaches to the classification of thermal portraits include the use of modern technologies of machine learning, image processing, and deep learning, which allow achieving high accuracy and speed in real time.

The use of machine learning and deep learning algorithms, in particular neural networks, is one of the most promising approaches to thermal portrait classification. Technologies such as CNN (Convolutional Neural Networks) allow the detection of complex patterns and specific thermal characteristics of objects that may be invisible to traditional image processing methods. By training on large datasets, neural networks can accurately classify different types of military objects based on their infrared thermal signatures. Machine learning, especially deep learning, has become a revolutionary approach in the field of image analysis, particularly thermal imaging. Thanks to their ability to detect complex patterns in data, neural networks such as CNNs have achieved significant success in classifying thermal portraits of military equipment.

So, the initial stage, and at the same time the most important in this work, is the creation and replenishment of the dataset, which involves searching for various images of fortifications from the sources that were already listed above. Thousands of video and photo materials were analyzed from sources such as social networks and pages with the corresponding to the image collection. It is worth noting right away that this part took the most time, approximately 90% of the work.

The main objects for localization in the dataset are mainly systems of trenches, dugouts, caponiers and trenches, which, as already noted, will be perceived as one single class, since they not only play a similar role in defense or are directly part of one fortification system, but also in the image frame it is difficult to recognize them as a separate type of field fortification.

As a result, the dataset, although it consists of virtually one class, its peculiarity lies in the fact that the images of fortifications are completely diverse. This set consists of more than a thousand images for training, more than 300 images for validation and about 40 images for testing. Although they were collected with the rule that the trench was photographed only from above, there are many other parameters that need to be taken into account.

Therefore, the following problems arose during the data collection process:

1. The environment of the fortification and its appearance;
2. Weather, place and time of shooting;
3. Specifics of uploading an image with a fortification;
4. Shooting quality and angle.

Unlike traditional image processing methods, where features are selected manually, neural networks learn to identify the most important features without human intervention. This allows them to detect subtle differences that may be invisible to the human eye.

Thermal images often contain complex nonlinear relationships between pixels. Deep neural networks are capable of modeling such relationships, which allows for high classification accuracy. Deep networks can be adapted to different types of data and tasks, making them a versatile tool for analyzing thermal images. CNNs are the

 ISBN 978-3-98924-119-0

most common type of neural network for computer vision tasks. They are particularly well suited for image analysis due to their ability to detect local features.

A large set of thermal images of various military objects is collected. Each image is annotated with the corresponding class (e.g., tank, BMP, artillery unit). A CNN architecture suitable for the task of thermal image classification is created. You can use ready-made models (e.g., ResNet, VGG) or develop your own. The model is trained on the dataset using the backpropagation algorithm. After training, the model is evaluated on a test dataset to determine its accuracy. The best model is integrated into the system for real-time use.

CNNs are capable of achieving high classification accuracy even under difficult conditions (e.g., in the presence of noise, interference, or small object sizes). CNNs can detect objects regardless of their size, rotation, or scale. The classification process can be fully automated, allowing large amounts of data to be processed. Large and diverse datasets are required to train effective models. Thermal images can vary greatly depending on weather conditions, time of day, and other factors. It is necessary to develop models that are resistant to enemy attacks, such as data substitution or the creation of misleading images. In the future, we can expect further development of deep learning methods, which will allow us to solve more complex tasks, such as tracking moving objects, determining their types, and evaluating their tactical and technical characteristics.

Thermal image classification systems can be integrated with other control and decision-making systems, allowing to create more efficient and intelligent systems. Deep learning, in particular CNN, is a powerful tool for classifying thermal images of military equipment. Thanks to its capabilities, this approach opens up new prospects for the development of surveillance and reconnaissance systems. However, to achieve even greater success, it is necessary to continue research in this area, in particular, to develop new neural network architectures, create large datasets, and develop methods of protection against attacks.

Choosing the optimal Convolutional Neural Network (CNN) architecture for classifying thermal images is a critical step in the development of computer vision

 ISBN 978-3-98924-119-0

systems, especially in areas such as the defense industry. Each architecture has its strengths and limitations that affect accuracy, training speed, and computational resources.

For small datasets, models with fewer parameters may be more effective to avoid overfitting. For complex tasks, such as object recognition in low visibility or in the presence of noise, more powerful models may be required. Limited computational resources may require the use of simpler models or model compression techniques. In real-world systems, image processing speed can be critical, requiring the use of lighter models.

Several convolutional neural network (CNN) architectures have become widely used in the classification of thermal images, each offering specific advantages depending on the complexity of the task and the available computational resources. Early architectures such as LeNet represent some of the first successful CNN designs. Although LeNet demonstrated high effectiveness in recognizing handwritten digits and other relatively simple patterns, its limited depth and representational capacity make it insufficient for more complex tasks such as the classification of thermal portraits, where subtle temperature variations and irregular object contours must be captured.

A significant breakthrough in deep learning came with AlexNet, one of the first deep neural networks to achieve outstanding results in the ImageNet competition. With a larger number of convolutional layers and filters, AlexNet is capable of extracting more complex and abstract features from images, making it far more suitable for thermal image analysis than earlier models. However, as the field advanced, researchers developed deeper and more structured architectures such as the VGG family, known for their uniform structure and considerable depth. Although VGG models provide improved accuracy in many computer vision tasks, they require substantial computational resources, which may limit their practicality for real-time thermal image processing.

More modern architectures, such as ResNet, introduced the concept of residual connections, allowing networks to be trained at unprecedented depths without encountering the vanishing gradient problem. For thermal image classification—where minute differences in texture and temperature distribution are critical—ResNet models often achieve high accuracy while maintaining training stability. A related advancement is the DenseNet architecture, which employs dense connectivity patterns so that information from earlier layers is reused throughout the network. This enables more efficient feature propagation and reduces the number of parameters, making DenseNet particularly effective in capturing detailed thermal patterns.

Other families of models focus on balancing performance with computational efficiency. The EfficientNet architecture scales depth, width, and resolution in a principled way, producing models that achieve strong performance even under resource constraints. This makes EfficientNet a suitable choice for systems that must process large volumes of thermal imagery while maintaining efficiency. In contexts where computational resources are limited even further—such as portable thermal imaging devices or embedded systems—MobileNet models offer an effective solution. Their lightweight design allows for real-time inference on mobile hardware, making them particularly advantageous for field applications where mobility and low power consumption are essential.

Together, these CNN architectures form a spectrum of solutions for thermal image classification tasks, enabling the selection of an optimal model depending on accuracy requirements, hardware limitations, and the complexity of thermal patterns characteristic of military targets.

The larger and more diverse the dataset, the more complex the model that can be used. Critical applications may require high accuracy, which necessitates the use of more complex models. If image processing speed is critical, lighter models should be used.

Using models pre-trained on large datasets (e.g., ImageNet) can significantly speed up the training process and improve results.

Increasing the number of training images using various transformations (rotation, scaling, noise) can improve the generalization ability of the model.

The use of regularization methods (L1/L2 regularization, dropout) can help avoid overfitting.

Table 1 - L1 and L2 regulatization comparison

L1 regulatization	L2 regulatization
Sum of absolute value of weights	Sum of square of weights
Sparse solution	Non-sparse solution
Multiple solutions	One solution
Buildt-in feature selection	No feature selection
Robust to outliers	Not robust to outliers

In the context of classifying thermal portraits of military targets, one of the key challenges in building reliable machine learning models is avoiding overfitting. Overfitting occurs when a model learns the training data too precisely, including random fluctuations, noise, and artifacts inherent in thermal imagery. Because thermal data often contains sensor noise, atmospheric distortions, and temperature anomalies, an overfitted model may incorrectly treat these distortions as meaningful features. As a result, such a model performs well on the training samples but fails to correctly classify new thermal images of real battlefield objects. This leads to high variance and poor generalization, which is critical in military tasks where systems must reliably distinguish between different types of targets—such as armored vehicles, infantry, or concealed equipment—under changing conditions.

Overfitting is especially common in non-linear and non-parametric algorithms, including deep neural networks used for thermal signature analysis. These models have a high degree of freedom, enabling them to construct overly complex decision boundaries that do not correspond to real physical characteristics of military targets. In the classification of thermal portraits, this may result in unrealistic models that misinterpret background noise, reflections, or sensor artifacts as structural features of targets.

Conversely, underfitting represents the opposite problem: the model is too simple to capture the essential thermal patterns of military objects. An underfitted system fails to learn the characteristic temperature distributions, silhouettes, and radiometric signatures that distinguish various military targets. As a result, it cannot accurately identify the dominant trends in thermal imagery, which is unacceptable for defense applications where early detection and correct classification directly influence operational decisions.

To address overfitting in the classification of thermal portraits, several techniques are commonly used. Regularization methods, such as L1 (lasso) and L2 (ridge) regularization, constrain model complexity by penalizing large coefficients. This prevents the model from overemphasizing minor variations in thermal data and helps it focus on stable, physically meaningful features of targets. L2 regularization gradually shrinks weight values toward zero while preserving their structure, whereas L1 regularization can push weights all the way to zero, effectively removing irrelevant or noisy features learned from thermal images.

Another widely used technique is dropout, which introduces stochasticity during training by randomly deactivating a portion of neurons in the neural network. This prevents the model from becoming overly dependent on specific feature combinations that might appear only in certain thermal conditions, such as a particular type of atmospheric noise or sensor distortion. By reducing the interdependence of network units, dropout encourages the development of more robust and generalizable representations of thermal targets. However, networks trained with dropout often require more training epochs to converge.

Together, these approaches help create classification models that generalize well across varied battlefield environments and sensor types. Preventing overfitting ensures that systems trained on thermal portraits remain reliable when encountering new operational scenarios, different weather conditions, or evolving military technologies. This robustness is essential for developing effective automated recognition systems that support surveillance, targeting, and threat assessment based on thermal imaging.

Choosing the optimal CNN architecture for thermal image classification is a challenging task that requires a deep understanding of both deep learning theory and data specifics. By carefully analyzing system requirements and experimenting with different architectures, high classification accuracy can be achieved.

 ISBN 978-3-98924-119-0

Table 2 - Comparison of convolutional neural network architectures

Architecture	Advantages	Disadvantages	Notes
LeNet	Simple, fast	Low accuracy for complex tasks	Suitable for small data sets
AlexNet	High accuracy	Requires a lot of computing resources	Classic model, but may be outdated
VGG	Deep, expressive	High computational costs	Good choice for large datasets
ResNet	Allows training of deep networks, high accuracy	Can be difficult to configure	One of the most popular architectures
DenseNet	Efficient use of information, high accuracy	May be more difficult to train than ResNet	-
EfficientNet	Optimal balance between accuracy and computational complexity	May be less effective for very complex tasks	Excellent choice for mobile devices
MobileNet	Very lightweight, fast	May have lower accuracy than more complex models	Ideal for embedded systems

Histogram of Oriented Gradients (HOG) is a powerful image processing method used to extract features, particularly useful for object recognition tasks. This method is based on the analysis of pixel brightness gradients, which allows describing the texture and shape of objects in an image.

First, the image is divided into small cells. For each pixel in a cell, a gradient is calculated that indicates the direction of the fastest change in brightness and its magnitude. For each cell, a histogram is created that describes the distribution of

 ISBN 978-3-98924-119-0

gradients by direction. Each bin of the histogram corresponds to a specific range of directions. To reduce the impact of changes in lighting and contrast, the histograms are normalized. The normalized histograms of the cells are combined into larger blocks, and a feature vector is calculated for each block. This vector is the object descriptor.

HOG methods are resistant to changes in lighting, which is an important property for thermal images, where brightness can vary greatly. HOG captures the textural characteristics of objects well, which is key to recognizing different types of military equipment. HOG descriptors are quite resistant to changes in scale, rotation, and minor deformations.

For objects with complex geometry or large internal variations, HOG may not provide sufficient accuracy. HOG can be sensitive to strong noise in the image, which can lead to misclassification. HOG works with grayscale images and does not use color information, which can be useful for some tasks.

Although HOG is a powerful method, it has its limitations. Modern deep learning methods, such as Convolutional Neural Networks (CNN), allow for the automatic learning of more complex and abstract image features.

HOG as the first layer: HOG descriptors can be used as input to CNNs, combining the advantages of traditional image processing methods and deep learning.

Using HOG for preprocessing: HOG can be used to preprocess images before feeding them into CNN, which can improve the quality of learning.

HOG is an important tool for image processing and can be effectively used for classifying thermal portraits. However, for best results, it is recommended to combine HOG with modern deep learning methods.

Based on the results of the initial research analysis, a flexible preprocessing pipeline was developed that allows for the automation of thermal image processing before further clustering. An important feature of this pipeline is the ability to flexibly adjust parameters at each stage, which allows the process to be adapted to different input data characteristics and segmentation requirements.

The first step in the pipeline is to load the image from disk. For this purpose, the standard cv2.imread function is used, which reads the image in BGR format. For further processing, the image is immediately converted to RGB format, since it is in this color space that it is more convenient to work with pixel values in the context of heat maps. In addition, an image in HSV format is additionally formed, which allows for further flexible selection of features for clustering.

One of the important problems when working with thermal images is the presence of noise caused by both the thermal imaging process itself and the image compression in JPG format. To reduce the impact of noise artifacts on the clustering results, the pipeline provides for the use of Gaussian blur. The blur level is set by the blur_level parameter and can be changed for experiments with different smoothing qualities. Using Gaussian blur allows you to suppress small-scale noise and emphasize large temperature regions in the image.

The next step is to crop the upper part of the image. This is an important step, since thermal imaging cameras often display service information (for example, temperature, time, sight marks) in the upper part of the frame. Saving this area leads to the appearance of artificial clusters during clustering, which do not carry any useful information. Therefore, the pipeline provides the crop_top parameter, which allows you to flexibly specify the number of rows for clipping.

After preprocessing, a set of features is formed for each pixel of the image. For this, a coordinate grid (x, y) is constructed, which allows you to take into account the spatial structure of the image during clustering. In addition to the coordinates of each pixel, the values of the selected color channels are added. An important feature of the implemented pipeline is the ability to flexibly choose which channels will be used for clustering. The 'R', 'G', 'B', 'H', 'S', 'V' channels are supported, which can be combined in any order. For example, for the analysis of the temperature structure, it may be sufficient to use only the Value channel, while for a more complete description of the image, it is advisable to use a combination of HSV or separate RGB channels.

To ensure the correct operation of clustering algorithms, all constructed features are subject to standardization using StandardScaler. This allows you to bring all features to a single scale, avoid the dominance of features with large numerical ranges (for example, x, y coordinates) over the values of color channels.

Thus, the proposed pipeline is quite universal and allows you to quickly prepare an image in the form of a feature matrix for further clustering. Due to the ability to variably adjust key parameters (blurring, cropping, channel selection), it allows you to conduct a deep analysis of the influence of each of these factors on the segmentation results.

The figure below shows an example of changing the data dimension after applying the preprocessing pipeline for different channel configurations.

One of the first models that was used for the thermal image clustering problem was the KMeans algorithm. This is one of the most popular and easy-to-implement clustering algorithms that allows you to split data into a given number of clusters based on the distance between pixels in the feature space.

In the implemented pipeline, the KMeans algorithm was applied to the feature matrix formed after image preprocessing. A feature of using this algorithm is the need to specify the number of clusters n_clusters before starting training. During the experiments, several runs of the algorithm were conducted with different numbers of clusters (mainly in the range from 3 to 6), which allowed us to evaluate the impact of this parameter on the quality of segmentation.

One of the important factors when using KMeans was also the selection of the optimal combination of features. As experiments showed, using only pixel coordinates (x, y) in combination with the green component (G) or HSV space allowed us to achieve a clearer separation of military equipment objects from the background. Adding all three RGB components often led to the formation of additional clusters in noise areas, so in most cases the optimal configuration was to use a smaller number of channels.

In terms of segmentation quality, KMeans demonstrated stable results on most images. By using Gaussian blur and cropping the upper part of the images, it was possible to minimize the influence of extraneous elements and noise. The algorithm distinguished the contours of the equipment well, however, in cases of a complex background or the presence of smooth temperature transitions, the clarity of the cluster boundaries was somewhat reduced.

The main advantage of KMeans is its high processing speed and ease of setup. At the same time, its disadvantages include sensitivity to the choice of the number of clusters and limited ability to model complex cluster shapes. However, in the tasks of primary segmentation of thermal images, this algorithm has proven itself as a basic and reliable tool.

The figures below show examples of the results of image segmentation using the KMeans algorithm with different numbers of clusters and different channel configurations. YOLO (You Only Look Once) is a revolutionary method in the field of image processing, especially for tasks that require fast detection and classification of objects in real time.

Unlike traditional methods that use sequential steps to detect and classify objects, YOLO processes the entire image at once. This allows for high speed performance. YOLO combines the tasks of detection and classification into a single regression model. This simplifies the architecture and increases efficiency. YOLO takes into account information from the entire image when classifying each object, which helps to avoid errors associated with isolated analysis of individual parts of the image. Thermal images usually have their own characteristics (e.g., low contrast, narrow range of values). Therefore, before being fed into the YOLO network, they can be further processed (normalization, filtering). A large dataset containing annotated images with marked objects and their classes is used to train YOLO on thermal images. During operation, YOLO divides the image into a grid and predicts the presence of an object, its class, and the coordinates of the bounding box for each grid cell.

Critically important for real-time systems such as surveillance systems. YOLO is capable of detecting objects of various sizes and shapes in thermal images. Relatively simple architecture compared to other methods. YOLO may have difficulty detecting very small objects. For objects with non-standard aspect ratios, accuracy may decrease. New versions of YOLO are constantly being released, improving accuracy and speed. YOLO can be combined with other methods, such as HOG, to improve accuracy.

YOLO models can be developed that are specifically adapted to work with thermal images, taking into account their characteristics. YOLO is a powerful tool for classifying thermal images, providing high speed and accuracy. However, like any other method, it has its limitations. The choice of a specific YOLO model depends on the specific task and available computing resources.

Modern computer vision methods actively use deep neural networks for object detection and classification. Among the most effective architectures in this field, YOLO is distinguished, which combines different approaches to image analysis.

Comparison of these models allows us to assess their advantages, limitations, and suitability for the tasks of automated recognition of plastic containers in vending machines.

The YOLO (You Only Look Once) model is one of the most common approaches in the field of computer vision for detecting objects in an image in real time. It combines high performance, compact architecture, and significant detection accuracy, which makes it an effective tool for applied tasks of object recognition and sorting — in particular, plastic containers in vending machines.

Unlike traditional approaches, where detection and classification are performed sequentially, YOLO implements a single-stage detection through the image, during which the following are simultaneously determined:

• the presence of objects;

• the coordinates of the frames (bounding boxes);

• object classes according to the trained set.

This approach is described in [14], where the authors proposed a unified architecture of a deep neural network that processes the entire image, forming spatial and semantic features in one step. Thanks to this, YOLO provides real-time detection, which is critically important for systems where decisions must be made instantly — for example, when feeding a bottle into a fandomat.

In the context of the plastic container sorting task, YOLO can be applied in two key aspects:

- Bottle detection in an image. This is the basic stage during which the system determines whether an object is present in the frame and localizes it. This allows it to filter out the background and extraneous objects.

 ISBN 978-3-98924-119-0

- Bottle parameter classification. Thanks to the multi-class classification capability, YOLO can not only identify a “bottle”, but also detect its characteristics — for example, transparency, the presence of a lid, or approximate volume. This approach provides a combined feature classification, which is particularly useful for automated sorting.

The advantage of the YOLO architecture is that the process is integrated into a single step, without the need for an external classifier. This reduces processing latency, which is a critical factor in real-time decision-making systems such as fandomats. According to a review [15], modern versions of YOLO (in particular, YOLOv5–YOLOv8) demonstrate a speed/accuracy ratio that exceeds most alternative methods such as Faster-RCNN or SSD, especially in real-time scenarios.

However, when using YOLO to classify combinations of parameters (e.g., “transparent/opaque”, “with/without a lid”, “small/medium/large volume”), the number of classes grows exponentially. For example, with three groups of parameters ($2 \times 2 \times 3$), the system must train on 12 classes, which makes data balancing and model maintenance difficult. This requires careful selection of the training sample and regular data updates as the packaging assortment changes.

From a technical perspective, YOLO supports exporting models to standardized formats (ONNX, TensorRT, CoreML), which enables integration with libraries such as OpenCV. This allows the model to be deployed even on low-cost devices with limited resources, such as embedded fan modules.

Innovative algorithms enable real-time processing of thermal portraits, which is critical for modern surveillance systems. This enables the instantaneous detection and classification of objects on the battlefield, thereby enhancing decision-making efficiency in combat conditions.

A thermal portrait (or thermogram) is an image that reflects the distribution of infrared radiation emitted by objects. Each object has its own unique thermal signature, which depends on its temperature, material, and size. This information can be used to identify various objects, even if they are camouflaged or in low-visibility conditions.

In combat situations, every second counts. The ability to instantly detect and classify objects allows operators to respond to changing conditions and make effective decisions quickly. Automatic processing of thermal images reduces false alarms and increases target detection accuracy. Thermal cameras can operate in low visibility, fog, smoke, and other obstacles, making them indispensable for surveillance in challenging weather conditions. Automatic processing frees the operator from routine work, allowing them to focus on more complex tasks.

Neural networks enable systems to be trained to recognize complex visual images, such as thermal portraits. They are capable of detecting objects of various sizes and shapes, even if they are partially obscured or camouflaged. Signal processing algorithms are used to filter noise, highlight characteristic features of thermal portraits, and increase their contrast. Combining data from multiple thermal cameras enables the creation of three-dimensional models of objects, thereby increasing the accuracy of their identification and tracking. Integrating data from various sensors (e.g., thermal, optical, radar) provides a more comprehensive picture of the situation and enhances the system's reliability.

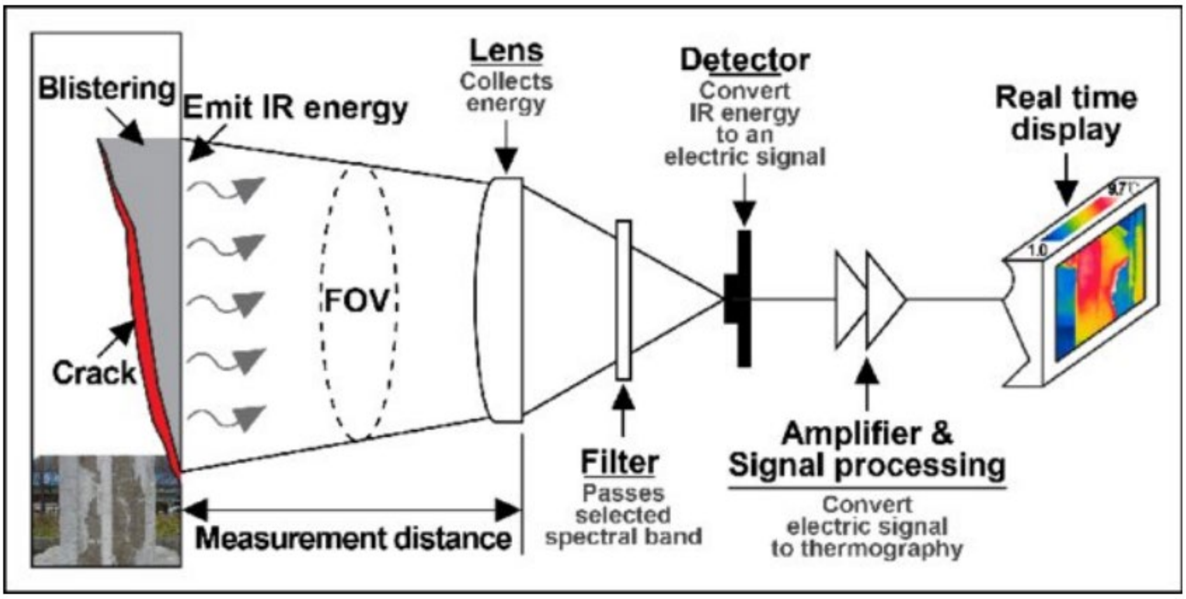

Fig. 3 - Schematic diagram of image transmission to a thermal camera

Modern computers and optimized algorithms enable the processing of large amounts of data in real-time. Many processes, such as object detection, classification, and tracking, can be automated, reducing the need for manual control. Thermal imaging processing systems can be easily integrated into various platforms, such as uncrewed aerial vehicles, ground robots, and surveillance systems. Thermal imaging, also known as thermography, is playing an increasingly important role in modern military technology. Its use has greatly expanded surveillance and reconnaissance capabilities, giving the military an advantage on the battlefield.

Thermal cameras detect infrared radiation emitted by all objects with a temperature above absolute zero. This radiation is focused on the camera's detector, which converts it into an electrical signal. This signal is then processed and transformed into the image we see on the screen.

Thermal portraits play a significant role in modern military operations, providing capabilities that are otherwise unattainable with conventional optical systems. In armored warfare, thermal sights installed on tanks enable crews to detect, identify, and engage targets even in complete darkness or under conditions of smoke, dust, and battlefield obscurants. The thermal contrast between objects allows tank operators to maintain situational awareness and conduct precision fire regardless of visibility.

Uncrewed aerial vehicles have also become essential platforms for thermal imaging. Equipped with high-resolution thermal cameras, drones perform reconnaissance, monitor enemy movements, and assist in fire adjustment. Their aerial perspective, combined with thermal sensing, makes it possible to detect hidden personnel, camouflaged equipment, and heat-emitting positions, significantly enhancing the effectiveness of intelligence gathering and target acquisition.

For infantry units, individual thermal night-vision devices provide critical support during nighttime operations. These devices allow soldiers to navigate terrain, detect opponents, and coordinate movements without relying on visible light sources that could reveal their position. By visualizing heat signatures, foot soldiers gain a tactical advantage in close-quarters engagements and in environments where traditional optics are ineffective.

Overall, thermal portraits significantly expand the operational capabilities of ground forces, armored units, and aerial reconnaissance systems, making them indispensable tools in contemporary military practice.

Thermal imaging technology is constantly evolving. New materials for detectors, more powerful processors, and machine learning algorithms are enabling the creation of increasingly sophisticated thermal cameras. In the future, we can expect increased resolution, allowing smaller objects to be detected at greater distances. Additionally, we can expect reduced size and weight of thermal cameras, making them more compact and lightweight, which will enable their use on a wider range of platforms. Thermal cameras will also be closely integrated with other fire control systems, navigation systems, and other military systems.

Thermal portraits play an important role in modern warfare. This technology allows the military to obtain critical information about the enemy and successfully complete their tasks. With the development of technology, thermal cameras will become even more effective and versatile tools in the hands of the military.

Thermal imaging has become an indispensable component of modern military operations, offering a wide range of capabilities that significantly enhance situational awareness, target detection, and operational effectiveness. One of its most important applications is the detection of camouflaged objects. Since any object with a temperature different from its surroundings stands out clearly on a thermogram, thermal cameras are able to reveal equipment, personnel, or structures even when they are visually concealed by camouflage, vegetation, or terrain features.

Night vision is another critical domain where thermal imaging proves essential. Unlike traditional night-vision devices that amplify visible or near-infrared light, thermal cameras operate independently of illumination and therefore allow continuous observation in complete darkness. This makes them indispensable for night raids, reconnaissance missions, and covert maneuvers.

Thermal imaging also plays a role in detecting mines and improvised explosive devices. Differences in soil temperature, caused by buried objects, can reveal thermal anomalies that point to hidden explosives. This capability increases the safety of engineering units and patrols operating in high-risk areas.

Effective target identification is another area where thermal imaging excels. By capturing subtle variations in heat signatures, thermal cameras allow operators to

 ISBN 978-3-98924-119-0

distinguish between infantry, armored vehicles, artillery systems, and other targets within seconds. Once detected, targets can also be tracked with high precision. Even in fog, smoke, or dust, moving objects maintain distinct thermal signatures, enabling reliable tracking and supporting coordinated fire control.

Beyond detection and targeting, thermal imaging assists orientation and navigation. Thermal maps of terrain created from aerial or ground-based sensors help troops navigate unfamiliar or obscured environments, reducing the risk of disorientation in complex landscapes or adverse conditions. Early detection of potential threats at long distances further enhances operational readiness by giving units time to react appropriately.

In military aviation, thermal imaging is widely used on night-attack aircraft and helicopters for locating and engaging ground targets during low-visibility missions. Similarly, in search-and-rescue operations, thermal cameras help locate missing or injured individuals on land or at sea by detecting their heat signatures even through vegetation or in poor weather.

A major advantage of thermal imaging systems is their all-weather capability. Unlike optical devices that degrade in fog, rain, or snow, thermal cameras maintain functionality under virtually all atmospheric conditions. Moreover, they operate passively—without emitting radiation—which makes them extremely difficult for the enemy to detect.

Modern thermal cameras also provide high-resolution imagery, enabling detailed analysis and supporting accurate decision-making. Their design allows easy integration into fire-control systems, surveillance networks, unmanned platforms, and other military technologies, making them a flexible and powerful component of contemporary combat systems.

Altogether, these capabilities demonstrate why thermal imaging has become a cornerstone technology in modern warfare, significantly enhancing both operational safety and combat effectiveness across numerous military domains.

Thermal imaging has become an integral part of modern combat equipment. Its capabilities are constantly expanding thanks to new technologies and image processing

 ISBN 978-3-98924-119-0

algorithms. The use of thermal imaging gives the military a significant tactical advantage on the battlefield.

Innovative algorithms for processing thermal images in real time are increasingly being used in various fields, including military, security, and industrial. These technologies improve the effectiveness of surveillance systems, ensure the safety of people and property, and open up new opportunities for the development of various industries.

Instead of using separate approaches to classify different types of objects, innovative systems use multi-task methods. They allow several tasks to be solved simultaneously, such as classification, localization, and prediction of object movement, which increases the efficiency and speed of thermal image processing.

Multi-task learning is an approach in machine learning that allows a model to simultaneously learn to perform several related tasks. Unlike traditional methods, where each task is solved by a separate model, the multi-task approach allows the model to extract common features from different tasks, which improves its generalization ability.

In the context of thermal image processing in the military sphere, a multi-task approach plays a key role, as it allows to significantly increase the efficiency of analyzing and interpreting thermal portraits of targets. Thanks to this approach, one model can simultaneously perform several interrelated tasks that previously required the use of separate algorithms. In particular, the models are able to simultaneously determine the type of object — for example, a tank, a vehicle or an infantryman — and localize its exact position in the image. The simultaneous solution of these two tasks allows for faster and more accurate detection and classification of targets in real time.

In addition, multi-task models can predict the further movement of objects, based on their previous trajectory, which is critically important for adjusting fire, assessing threats or optimizing the route of movement of military units. No less important is the ability of such models to detect anomalies — objects or situations that differ from the typical picture. These can be unknown thermal signatures, unusual movement patterns, or potentially dangerous anomalies that signal the presence of a hidden threat.

 ISBN 978-3-98924-119-0

The advantages of the multi-task approach are that joint training on different types of tasks allows the model to form more informative and discriminative features that better describe the thermal portraits of objects. This improves the quality of classification and localization accuracy even in difficult conditions, such as noise or low contrast images. In addition, instead of having to train and run several separate models for different tasks, one integrated multi-task model is enough, which reduces computational costs and the need for large amounts of data for each task separately. This approach makes the system more flexible, compact, and suitable for use in real combat conditions, where the speed of decision-making is a decisive factor.

Different tasks use common initial layers to extract common features and then branch out into separate branches to perform their tasks. The loss function of the model is the sum of the losses for each individual task. The attention mechanism allows the model to focus on different parts of the input data to perform different tasks.

Attention mechanisms, particularly multi-head attention, play a critical role in multifunctional thermal image processing tasks, as they allow the model to simultaneously analyze different aspects of the scene and adaptively allocate computational resources between multiple tasks. In the context of thermal analytics, such mechanisms enable the model to “focus” on the most informative areas of the image, taking into account the intensity of the thermal signal, the contrast between objects and the background environment, local heating patterns, and the dynamics of object movement. This is especially important when the model has to perform several functions simultaneously—for example, detecting objects, classifying them, analyzing thermal characteristics, and predicting behavior.

Various neural network (Fig. 4) architectures are used to implement multi-task learning.

In the context of classifying thermal portraits of military targets, the multi-head attention mechanism provides a powerful way to extract diverse and complementary features from thermal data streams. Instead of relying on a single attention pattern, the model forms several parallel “views” of the same scene. Each attention head reacts to different characteristics of the thermal image: one may emphasize temperature

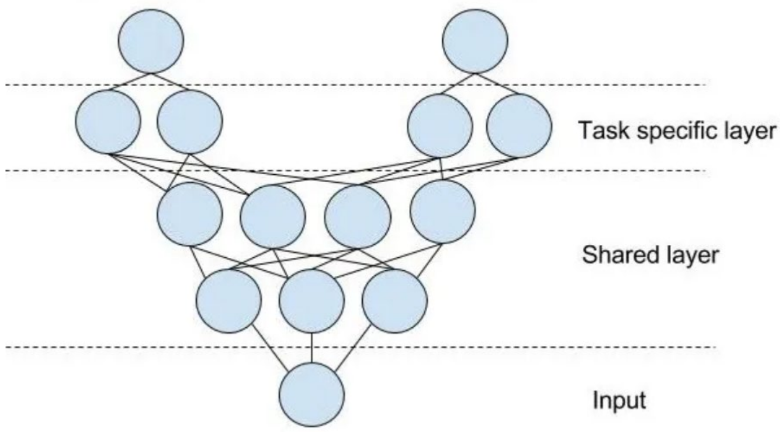

Fig. 4 - Visualization of a neural network

gradients that reveal object boundaries, another may isolate structural contours typical of military equipment, a third may track temporal fluctuations of heat, while a fourth analyzes spatial relationships between objects in the field of view. This creates a multilayered representation that significantly improves recognition accuracy in complex operational environments.

When the system must differentiate between various categories of military targets—such as personnel, vehicles, unmanned platforms, or heated infrastructure—multi-head attention helps separate these thermal signatures more effectively. Individual heads can focus on characteristic patterns, such as the stable thermal emission of a human body, the distinctive heat distribution of an engine block, or the cold geometric forms that serve as background context. This parallel extraction of features allows the model to localize targets precisely while maintaining sensitivity to class-specific traits.

In scenarios where both detection and classification must occur simultaneously, the attention mechanism distributes computational focus between regions responsible for identifying targets and those containing information essential for distinguishing their type. This prevents one task from suppressing the other—for instance, ensuring that detection performance does not mask the subtle cues needed for accurate classification of similar thermal silhouettes.

For broader tasks such as battlefield situational awareness, recognition of hostile assets, and navigation of autonomous systems under limited visibility, multi-head attention becomes even more critical. Some attention heads analyze fine-grained thermal details for identifying targets, others interpret global spatial geometry for orientation, and additional heads extract contextual cues such as environmental temperature, atmospheric distortions, or background thermal noise. This layered processing supports reliable operation even in challenging conditions with high variability of signatures, camouflage, or clutter.

As illustrated in Figure 5, the multi-head attention mechanism provides the architectural foundation for building flexible, resilient, and multitask-capable models used in the analysis and classification of thermal portraits of military targets. By processing multiple attention heads in parallel, the system can capture a wide range of thermal features—from fine-grained local patterns to broad contextual relationships—which strengthens robustness to noise and enhances performance under rapidly changing temperature or environmental conditions. This diversity of attention pathways directly contributes to maintaining high classification accuracy across complex operational scenarios.

Achieving optimal results with this mechanism, however, requires careful balancing between competing tasks such as detection, segmentation, and classification, ensuring that no single task dominates the learned features. The choice of neural network architecture must therefore reflect both the mission requirements and the structure of available thermal data. Finally, interpretability remains essential: understanding how the model distributes attention across heads, and how it arrives at specific decisions, is critical for assessing reliability and establishing trust in automated target recognition systems deployed in real-world military environments.

A multi-task approach is a powerful tool for processing thermal images. It allows for increased efficiency and accuracy of systems using thermal cameras in various applications.

Modern approaches to thermal portrait classification also use combined processing of data from various sources—thermal cameras, radars, drones, satellites, etc. Integrating such data improves classification accuracy by reducing the likelihood

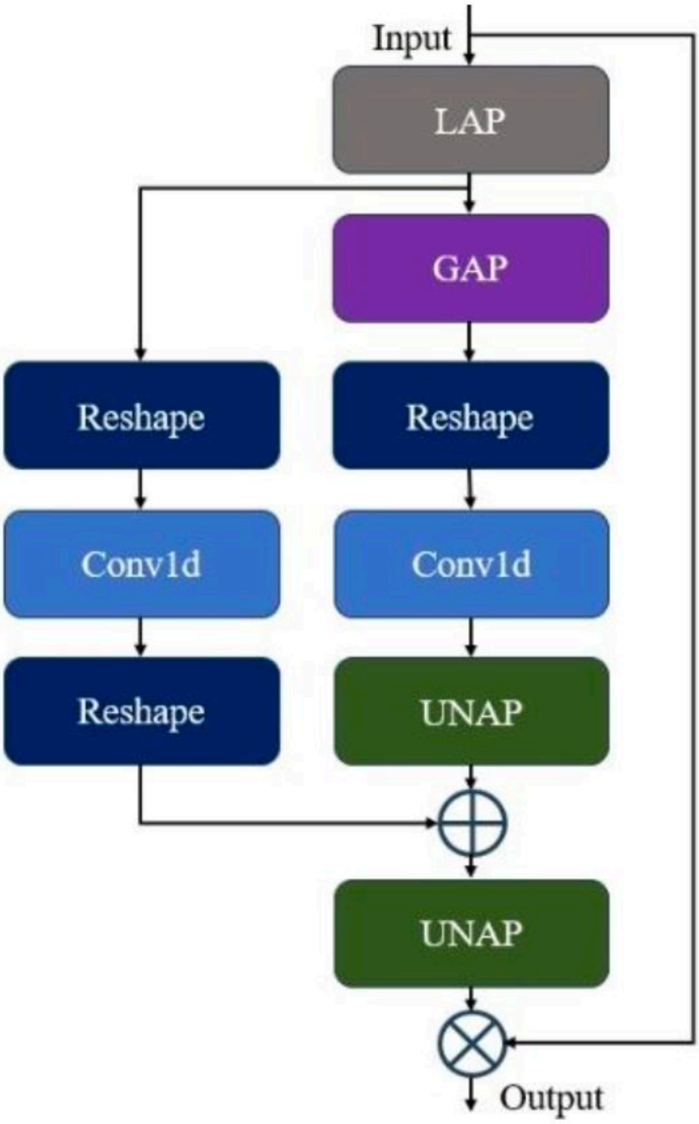

Fig. 5 - Visualization of a neural network

of errors caused by the limited capabilities of individual sensors.

Combining data from various sources, such as thermal cameras, radars, drones, and satellites, is one of the most promising areas in modern thermal portrait classification. This approach allows for a more complete and detailed picture of the object, which significantly increases the accuracy and reliability of classification.

Different sensors provide different information about an object. For example, thermal cameras capture thermal signatures, radars capture reflected radio waves, and optical cameras capture visible light. By combining this data, it is possible to obtain a more complete picture of the object and its characteristics. Each sensor has its limitations. For example, thermal cameras can be sensitive to weather conditions, and radars to interference. By combining data from different sensors, it is possible to reduce the impact of these limitations and increase the reliability of classification. Combining data allows us to discover new features of objects that cannot be detected by individual sensors. For example, combining data from a thermal camera and radar can allow us to determine the type of material an object is made of.

There are several approaches to combined data processing. Data from different sensors is combined at an early stage of processing, for example, at the raw signal level. This allows a single representation of the object to be obtained, which is then used for classification. In addition to early fusion, there is also late fusion. Data from different sensors is processed separately, and then the results are combined at the decision level. For example, classification results obtained using a thermal camera and radar can be combined using rules or neural networks. There are also hybrid approaches that combine elements of early and late fusion.

As illustrated in Figure 6, which presents a general example of a recurrent neural network (RNN) approach, the combined processing of data originating from different sensor systems relies on a broad set of algorithms designed to produce coherent and informative results. Central to such systems are advanced neural architectures—primarily convolutional neural networks (CNNs) and recurrent neural networks (RNNs). These models demonstrate high effectiveness in handling large and heterogeneous data streams, including thermal imagery, video sequences, and telemetry from multiple sensor modalities.

Their strength lies in the ability to automatically extract meaningful features and capture both spatial (via CNNs) and temporal (via RNNs) dependencies. This capability is particularly important for analyzing dynamic thermal portraits of military targets, where temperature distributions, motion patterns, and contextual variations must be interpreted jointly. By integrating data of different nature into a unified representation, these architectures enable more accurate, stable, and context-aware classification within complex operational environments.

 ISBN 978-3-98924-119-0

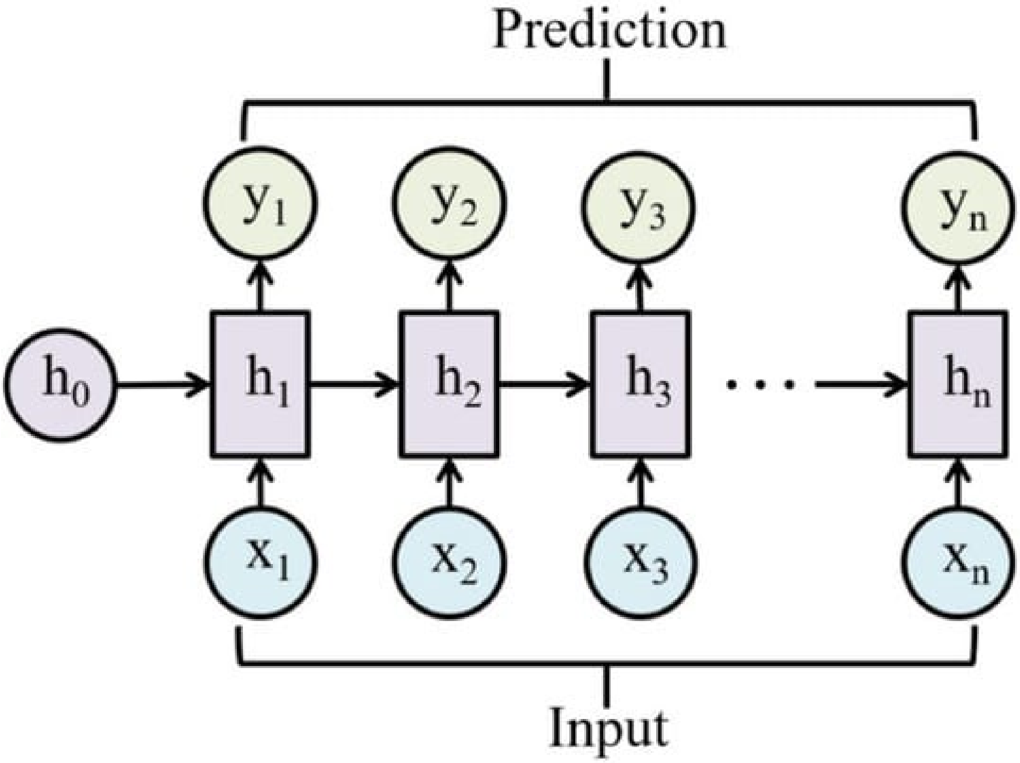

Fig. 6 - General recurrent neural network architecture solution

In addition to neural networks, classic machine learning methods, such as classification, regression, and clustering algorithms, play an important role. They can be used to reconcile results obtained from different observation systems or to build generalized models that identify common patterns in the data. Due to their interpretability, these methods enable the understanding of the structure of objects and the relationships between indicators obtained from sensor channels.

No less important are approaches based on decision theory, such as evidence theory or probabilistic models. Their application allows you to correctly combine vague, incomplete or contradictory data, which often arise during field observations. Such methods allow you to assess the degree of reliability of information, reconcile different sources and form the most justified generalized decision. Thanks to this, combined data processing becomes more stable, reliable and adaptive to real operating conditions.

By using additional information obtained from different sensors, higher classification accuracy can be achieved. Combined data processing reduces the impact of errors associated with individual sensors. Data combination allows more complex tasks to be solved, such as object recognition in conditions of interference and masking.

Data from different sensors can be obtained at different times and have different resolutions. This complicates their integration. The choice of the optimal data integration method depends on the specific task and available resources. Interpreting the results obtained by integrating data from different sensors can be difficult.

 ISBN 978-3-98924-119-0

CHAPTER 2
ONTOLOGY-BASED APPROACHES FOR CLASSIFYING MILITARY EQUIPMENT

Combined processing of data from different sources is a promising direction for the development of thermal portrait classification systems. This approach improves the accuracy, reliability, and capabilities of such systems, making them indispensable in many areas, from military affairs to security and environmental monitoring.

The use of ontologies for thermal portrait classification is a new approach that allows structuring knowledge about different types of targets. Ontologies help classification systems understand the relationships between different objects, their thermal properties, and possible interactions with other systems. This improves classification accuracy in rapidly changing environments.

An ontology is a formal model that describes the conceptual structure of a subject area. In the context of thermal portrait classification, ontology represents knowledge about different types of objects, their characteristics (size, shape, thermal signature), and the relationships between them.

Ontologies provide a clear structure for knowledge about objects, making it easier to organize and search for them. Thanks to ontologies, the system can understand not only the individual features of an object, but also their semantics and relationships. Ontologies can be easily updated and expanded to accommodate new types of objects or changes in their characteristics. Thanks to a deeper understanding of the subject area, classification systems that use ontologies can achieve higher accuracy.

First, an ontology is created that describes all possible types of objects, their attributes, and relationships. Thermal portraits are annotated according to the terms of the ontology. This allows each pixel of the image to be associated with a specific object class or part of it. Based on the annotated data, a classification model is trained that uses ontology to determine the class of an object in a new image.

The figure shows a simplified ontology for classifying military equipment. It includes classes such as "tank," "armored personnel carrier," "helicopter," and their subclasses. Each class has its own attributes, such as size, shape, engine type, and thermal signature.

Ontologies can be integrated with various technological approaches to improve the efficiency of thermal portrait classification, forming a more structured and substantiated data processing model. In combination with deep learning methods, ontological structures help build interpretable models in which knowledge about object types, their characteristics and relationships complement automatically obtained features. This provides better explainability of the work of neural networks and increases the reliability of results.

In the context of building knowledge bases, ontologies allow you to systematically expand information about objects, their behavior and properties, which is especially important for the analysis of thermal data in dynamic conditions. They provide a standardized representation of information that can be used to improve the processes of generalization, search and logical inference.

At the same time, ontologies play a key role in the development of expert systems, as they form a conceptual framework on the basis of which such systems can draw conclusions even when the data is incomplete, contradictory or ambiguous. Thanks to this, expert systems working with thermal portraits are able to make informed decisions in complex operational conditions, ensuring increased accuracy and stability of work.

An example of ontology for the classification of military equipment is shown in Fig. 7.

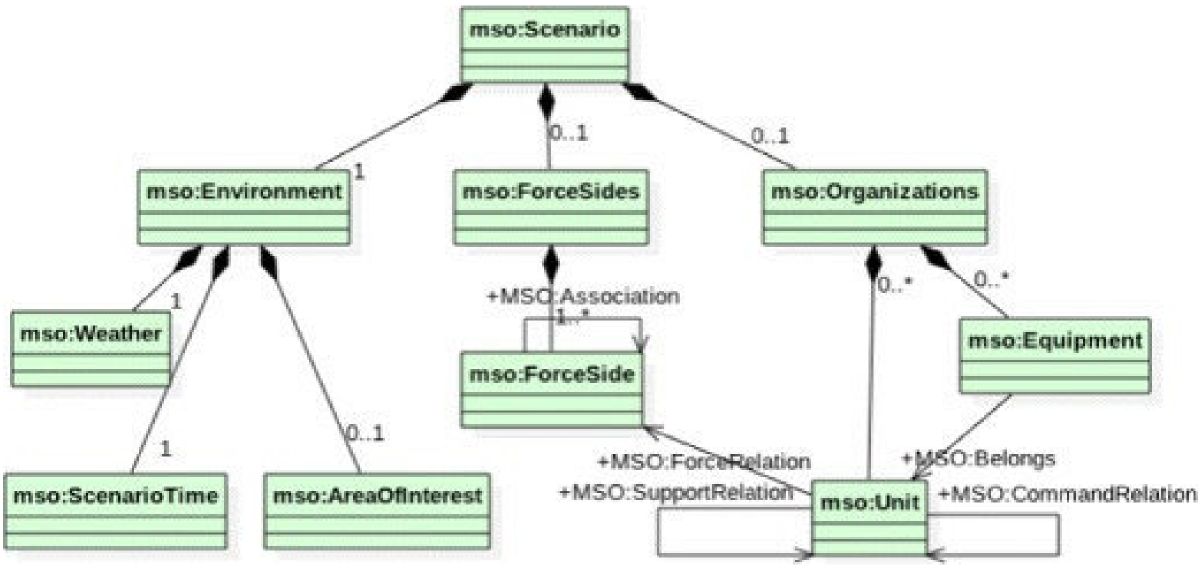

Fig. 7 - Example of an ontology for classifying military equipment

Ontological knowledge bases are a powerful tool for organizing and processing large amounts of information. They allow knowledge to be formalized in a clear structure, where each element of the base has a clearly defined meaning, properties, and relationships with other elements. They are particularly useful for working with data that has a complex hierarchical or relational structure, as is the case with the analysis of thermal portraits of targets.

Thermal portraits of targets obtained through thermal imaging systems contain a large number of characteristics, such as surface temperature, dimensions, shape, and other details. An ontological knowledge base allows you to organize this data into a clear hierarchical structure, where each object (for example, the type of equipment, its characteristics) will have a clear definition. This allows objects to be classified by type, similar or identical objects to be recognized, and rules to be created for their further analysis.

Thanks to ontological knowledge bases, it is possible to automate complex decision-making processes related to the classification and analysis of thermal portraits. The system can not only process input data, but also make decisions based on previously created rules or models, which reduces dependence on the human factor. This is critically important in combat situations, where every second counts.

An ontological knowledge base allows data from different sources to be integrated and analyzed in the context of existing models and patterns. For example, if the system receives a thermal image of an object, it can automatically compare this image with a database of equipment and determine the type of object, taking into account additional parameters (temperature, dimensions). This allows for quick and accurate identification and classification of the target, reducing the likelihood of errors that may arise due to human uncertainty or inaccuracies in observation.

The use of ontological knowledge bases significantly reduces the likelihood of human error. Decisions made based on automated rules are more stable and less prone to error than those based on the operator's intuitive perception. Military operations require accurate and timely information to minimize risks, and ontological knowledge bases can provide this.

ISBN 978-3-98924-119-0

Ontologies are easy to update and supplement with new data or rules, making them adaptable to changes in technology or situations. If new types of equipment appear or conditions on the battlefield change, the system can be quickly reconfigured to work with new data without the need for radical changes to the structure.

In combat conditions, a system using thermal imaging cameras can use ontology to determine whether an object is an enemy tank, armored personnel carrier, or other type of equipment. Thanks to ontology, it can collect all the necessary information (size, temperature, shape, movement) and automatically assign a category to it.

If the system detects a change in temperature or movement on an object, it can immediately review existing rules and adapt its response. For example, if an object changes location or shape, the system will automatically compare these changes with known patterns in the database to refine or confirm the identification.

Thanks to ontology, the system does not simply classify objects, but understands their nature, functions, and place in the overall picture. The use of machine learning methods allows many processes to be automated, such as object detection, characterization, and behavior prediction. The system is capable of effectively working with large amounts of data coming from various sources, ensuring prompt decision-making . By analyzing past data, the system can predict the future actions of objects, allowing it to anticipate developments on the battlefield.

The ontological knowledge base contains information about various types of military equipment, their characteristics (size, shape, thermal signature), types of weapons, possible actions, and the relationships between them. It classifies objects by type and subtype, determines the relationships between objects, and provides context for understanding the situation. Thermal images are collected from various sources (drones, satellites, ground sensors). Images are filtered, normalized, and segmented. Neural networks are used to detect objects in thermal images. The type of object is determined based on its characteristics and ontological information. Recurrent neural networks will be used to predict the movement of objects.

The use of ontologies for the classification of thermal portraits is a promising area of research. It allows the creation of more intelligent and adaptive systems that are capable of working effectively in complex conditions.

In order to improve the quality of thermal images, innovative methods of noise reduction and image sharpening are used. It is important to preserve important information about thermal signatures while reducing the impact of noise, which allows for more accurate identification even in conditions of poor visibility or external factors.

The quality of thermal images significantly affects the effectiveness of the systems that use them, especially in the fields of security, defense, and industrial control. Noise and low contrast can complicate the detection and identification of objects. Therefore, the use of image processing methods to improve them is critically important.

Noise in thermal images can arise for various reasons and significantly affect the quality of thermal portraits, and therefore the accuracy of subsequent classification and analysis. One of the key factors is atmospheric conditions: fog, rain, snow or smoke create additional distortions that worsen the contrast and reduce the visibility of objects in the thermogram. Such conditions lead to scattering of infrared radiation, which reduces the clarity of the thermal image.

An important source of noise is also the characteristics of the detector itself. Any thermal imaging sensor has its own level of internal noise, which arises from thermal fluctuations in the detector material, as well as quantization noise that appears during digital signal processing. These factors can cause uneven brightness and the appearance of artifacts that complicate the analysis of the thermal scene.

A separate problem is the limitations of the dynamic range. If the scene simultaneously contains very hot and very cold objects, the detector may not be able to correctly display all temperature values. As a result, information about individual areas of the image is lost or distorted, which reduces the informativeness of the resulting thermal portrait and complicates the correct selection and classification of targets.

Methods for reducing noise and improving image quality include various filtering and deconvolution techniques (example on Figure 8), as well as methods based on artificial neural networks. Filtering can be performed using spatial filters, such as

 ISBN 978-3-98924-119-0

average, median, or weighted average filters, as well as frequency filters, among which Wiener and Kalman filters are popular. Bilateral filters and NL-means filters are used for nonlinear filtering. Deconvolution aims to reduce blurring caused by optical system limitations or object motion.

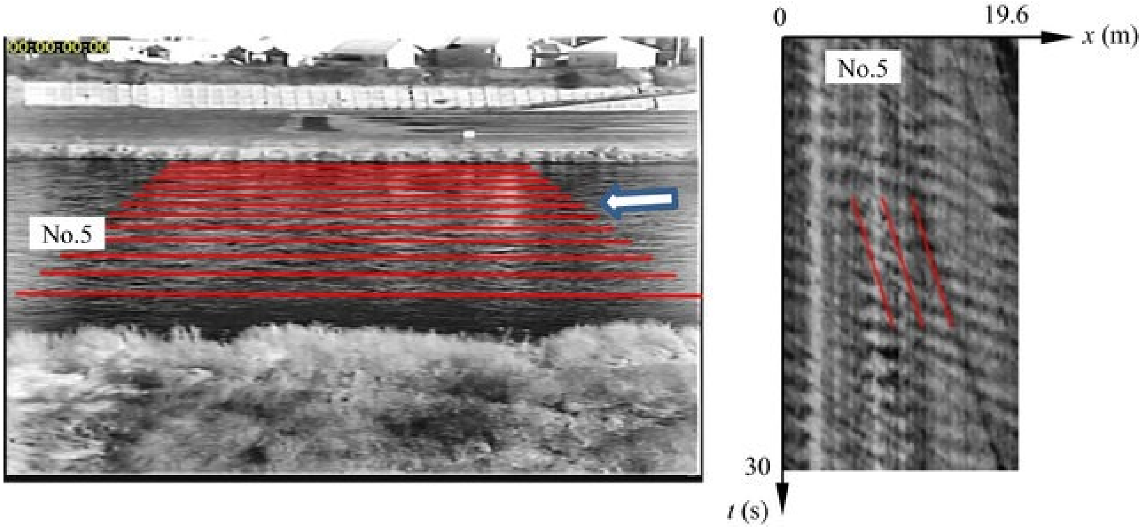

Fig. 8 - Noise in thermal images

The application of artificial neural network methods encompasses a range of architectures, including autoencoders for image restoration and noise suppression, generative adversarial networks (GANs) for synthesizing realistic thermal imagery, and U-Net models for segmentation and object detection. These approaches are particularly valuable for enhancing the quality and informativeness of thermal portraits of military targets, enabling more accurate classification and recognition under challenging operational conditions. The outcomes of these methods, including examples of restored, generated, and segmented thermal images, are visualized in Figure 9, which illustrates their effectiveness in practical scenarios.

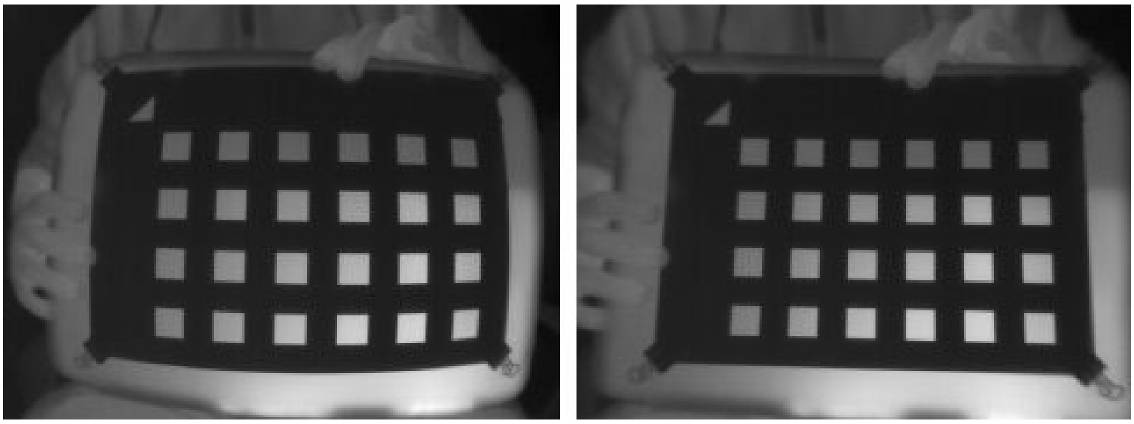

Fig. 9 - Comparison of a thermal infrared image before and after correction

ISBN 978-3-98924-119-0

Contrast enhancement methods include several techniques that improve image detail. Histogram equalization is used to expand the dynamic range of an image, improving the visibility of details in the background. Local adaptive histogram equalization allows contrast to be enhanced in specific areas of the image, which is particularly useful for images with uneven lighting. Reconstruction methods, such as sparse decomposition, help to represent the image as the sum of a small number of basic elements, which preserves important information during processing. An important part of such methods is the preservation of thermal fingerprint information. This includes accurately determining the temperature of objects, which is critical for identifying materials and the condition of objects, analyzing temperature distribution, which can indicate the presence of heat sources or moving parts, and accurately determining the shape and size of objects, which is necessary for their identification.

The choice of the optimal image processing method depends on the specific task, the characteristics of the thermal camera, and the level of noise in the image. Modern methods based on artificial intelligence demonstrate high efficiency in noise removal and improvement of thermal image quality, which allows for increased accuracy of automatic identification and analysis systems.

Systems that use AI are capable of adaptive learning based on new data. This allows classifiers to continuously improve their models and increase classification accuracy based on new types of targets appearing on the battlefield. Innovative methods use active learning approaches, where the system can independently request additional labels to improve models.

Adaptive learning means that the AI system can constantly update its models based on new data it gets while it's running. This is super important in military situations, where things on the battlefield are always changing, with new weapons and tactics popping up. The models always match the latest changes in the environment. Continuous retraining improves classification accuracy. New data helps identify and correct errors in models. The system becomes more versatile, capable of recognizing new types of targets.

Active learning is a method in which the system itself determines what data it needs for further learning. Instead of learning from all available data, the system focuses on the most informative examples that will help it improve its capabilities faster and more efficiently.

The system begins training on a small set of manually labeled data. The system selects unlabeled data that it believes is most important for improving the model (for example, data that is on the border between classes). An expert labels the selected data. The new labeled data is added to the training sample. The model is retrained on the expanded sample.

The combination of adaptive and active learning creates a powerful tool for developing AI systems that are constantly improving and adapting to new conditions. The system can independently determine what data it needs to improve its capabilities and then use that data to update its models.

Creating algorithms that more effectively select informative data. Combining active learning with deep learning, transfer learning, and reinforcement learning technologies. Creating convenient tools for experts that allow them to interact effectively with the system and provide the necessary labels.

Adaptive and active learning open up new opportunities for developing AI systems that are capable of continuous improvement and adaptation to changing conditions. These technologies have great potential for application in military systems that use thermal portraits, allowing for increased efficiency and reliability.

These innovative approaches to thermal portrait classification reduce the risk of false positives, increase target detection efficiency, and improve decision-making and interaction between different systems on the battlefield.

The development and classification of thermal portraits of military targets is a crucial area of innovative technology in the fields of security and defense. In this context, the use of modern thermal imaging systems significantly enhances the effectiveness of surveillance and object detection in conditions of limited visibility, such as at night, in smoke, or in fog. Thermal imaging camera technologies and machine learning algorithms, particularly deep learning and neural networks, enable the accurate identification of objects based on their thermal characteristics.

 ISBN 978-3-98924-119-0

In particular, the introduction of ontological approaches to the classification of thermal portraits enables the structuring of knowledge about objects and enhances the accuracy of analysis. The use of machine learning methods, such as CNN, significantly enhances the ability to automatically classify and identify objects, even in challenging conditions. At the same time, noise reduction and image quality improvement methods allow for the elimination of interference caused by atmospheric factors or detector characteristics.

Integrating data from various sources, such as thermal cameras, radars, and satellites, significantly improves the accuracy and reliability of classification, providing a more complete picture of an object or situation. An important advantage is the multitasking approach, which enables the simultaneous performance of multiple tasks, such as classification, localization, and prediction of object movement.

The use of artificial intelligence for adaptive learning contributes to the continuous improvement of models, allowing systems to respond quickly to new challenges on the battlefield. This enhances the versatility and accuracy of classification, which is particularly crucial for real-time military operations.

In modern military operations conducted both during the day and at night, a crucial component is the ability to detect a potential enemy first using automated surveillance systems. This is achieved through the use of thermal imaging cameras, which allow information to be obtained in conditions of poor visibility, particularly at night or in difficult weather conditions. Thanks to improvements in thermal imaging technologies and the use of effective video stream processing algorithms, the quality of images provided to the operator for analysis and decision-making has been significantly improved.

There are many methods for processing video streams, including digital filters that help reduce the effects of noise, blurring, and other distortions while increasing the contrast and clarity of images. This makes it possible to improve the accuracy of object detection in video streams from thermal imaging cameras. However, despite the development of numerous methods and techniques for improving the quality of video streams, there is no general theory of thermal imaging processing. Each method of

 ISBN 978-3-98924-119-0

processing video for further analysis must be evaluated through the prism of the observer's subjective perception. Visual assessment of video quality is a highly subjective process, which makes the definition of a "high-quality video stream" a somewhat vague benchmark on which to evaluate the effectiveness of the algorithms used.

The development of an algorithm for the preliminary processing of video streams obtained by thermal imaging cameras is an important step in the processing of thermographic data. Thermal imaging cameras allow you to obtain an image that reflects the temperature fields of objects in different conditions, which is critically important for applications in military technology, security systems, search and rescue, etc.

Infrared thermal imaging cameras are important tools in the field of surveillance and object detection due to their ability to detect thermal radiation from objects. These cameras specialize in recording infrared radiation, which is invisible to the human eye but allows images of the thermal characteristics of the environment to be obtained. They are capable of visualizing temperature differences on surfaces, making it possible to detect objects that emit heat, regardless of the lighting conditions of the environment.

The main advantage of infrared thermal imaging cameras is their ability to operate in conditions where traditional optical systems based on visible light cannot be effective. This includes operation in low visibility conditions, such as fog, smoke, rain, or at night. In such conditions, conventional cameras using visible light cannot provide clear information, while thermal imaging cameras effectively capture thermal radiation, allowing objects to be distinguished.

Military operations usually take place in low visibility conditions, especially at night or in difficult atmospheric conditions, when traditional optical devices become useless. Infrared thermal imaging cameras allow you to observe objects that may be invisible or difficult to identify with conventional cameras. This makes it possible to detect military targets, equipment, and potential threats that do not have clear contours visible under normal lighting conditions.

Thermal imaging cameras are also used to detect temperature anomalies in the environment, allowing, for example, hidden objects that emit heat to be detected, or the environment to be scanned for the presence of living organisms such as humans or animals. This is made possible by the cameras' sensitivity to temperature differences, which allows for quick assessment of the situation even in difficult conditions.

Thus, infrared thermal imaging cameras are indispensable tools for performing surveillance and target detection tasks in real time when other surveillance methods become ineffective, which is especially critical for the successful conduct of military operations.

The main steps in developing an algorithm for pre-processing video streams include the following stages:

The process of collecting and preparing data for processing video streams from thermal imaging cameras is a critically important stage, since its quality determines the effectiveness of further work with the data and the accuracy of the results obtained during analysis.

At the data collection stage, thermal imaging cameras are used to capture video from objects in different temperature conditions. This can include both stationary and moving objects. To ensure the reliability and accuracy of the results, video streams are collected in different conditions: during the day and at night, in different weather conditions, at different distances from the camera. Field trips allow you to capture real-life scenarios that reflect the different variations in conditions that thermal imaging systems may encounter.

After collecting video streams, an important step is their pre-processing, which includes noise removal, contrast correction, brightness and temperature normalization, as well as correction of errors that may arise due to bad weather conditions or problems with the technical condition of the camera. To ensure correct video processing, filtering or frame correction can be used to improve image quality and accuracy of results.

Data preparation involves converting the received video streams into a format that is convenient for further work, allowing for automatic analysis. Important aspects of preparation include vectorization, highlighting key objects and scenes, as well as normalizing temperature data to compare different video streams and ensure their convenient further use.

Data collection and preparation are necessary to ensure high-quality end results. Incorrect or insufficiently thorough processing can lead to inadequate conclusions, low accuracy in object recognition, or other errors in the system's operation. Field testing allows for the collection of real and diverse data, which helps to identify potential problems in processing video streams in real-world conditions.

Thermal imaging cameras often store video in different formats. Each format has its own characteristics that affect the efficiency of data processing:

• .avi (Audio Video Interleave): This format is one of the most common for storing video. It supports high-quality video and can contain a large number of frames. However, for thermal imaging, it is important to note that this format may not be the most efficient for storing large amounts of high-resolution data.

• .mp4 (MPEG-4 Part 14): The MP4 format is popular due to its high compression ratio, which allows for a significant reduction in video file size without significant loss of quality. This format is convenient for working with real-time video streams, as it reduces storage and network bandwidth requirements.

• .mkv (Matroska Video): The MKV format can contain multiple video and audio tracks, as well as subtitles and other additional data. It allows to store larger files with high quality, which can be useful when processing high-resolution video from thermal imaging cameras.

When choosing a format for storing video from thermal imaging cameras, it is important to consider not only the convenience of data storage, but also its real-time processing. Various libraries and frameworks can be used for further work with such videos, for example, OpenCV, FFmpeg, or GStreamer.

Thermal imaging cameras are usually capable of shooting video at high frame rates — from several tens to several hundreds of frames per second. High frame rates are important for tracking moving objects and accurately determining temperature changes in real time.

High frame rates in thermal imaging allow for smoother, more detailed video sequences, but they come with some challenges. One such challenge is the uneven movement of objects: due to the high frame rate, objects in the image may appear shifted or distorted, making it difficult to track them accurately. In addition, insufficient synchronization or problems with frame formation can lead to a blurry display in real-time, which reduces the accuracy of measuring the temperature of objects.

To solve these problems, frame interpolation is used, which allows you to add intermediate frames between existing ones. There are two main interpolation methods: linear and bicubic. Linear interpolation involves evenly distributing intermediate frames between two adjacent ones and works well when the changes between frames occur gradually. This method is simple and fast, but may not be effective enough for complex video streams with abrupt changes. Bicubic interpolation provides higher frame-to-frame detail accuracy, creating a smoother, more natural-looking image, but it requires significantly more computing resources. Thus, the choice of interpolation method depends on the desired balance between frame-to-frame accuracy and processing speed.

For thermal imaging cameras, it is important that the interpolation takes into account not only geometric parameters (the position of objects), but also temperature changes. Therefore, during interpolation, special attention should be paid to preserving temperature values and preventing them from being "stretched" or distorted during transitions between frames.

Frame alignment is an important part of preparing data for further processing, especially if video streams from a thermal imaging camera contain various types of distortions or object movements. Alignment helps synchronize the video and ensure that all frames are correctly matched, which is essential for further analysis. The primary alignment methods are optical alignment and stereo alignment technologies. Optical alignment: This technique utilizes the detection and tracking of points in images to align the video stream. Various feature detection algorithms can be used for this, such as SIFT (Scale-Invariant Feature Transform) or ORB (Oriented FAST and Rotated BRIEF). They allow you to find common points in different frames and align them to achieve a correct image. Stereo alignment technologies: If multiple thermal imaging cameras are used for video processing, stereo alignment may be required to synchronize frames obtained from different viewing angles.

Data collection and preparation are the foundation for effective processing of thermal camera video streams. Choosing the video format, interpolating, and aligning frames improves the quality and accuracy of the images obtained, which is essential for further analysis stages such as segmentation, object detection, and temperature measurement.

Video obtained from a thermal imager may contain various types of noise, including electronic noise, variable lighting, or object movement. To eliminate them, filtering methods such as median, Gaussian filters, or low-frequency filtering are used. A median filter is used to remove sporadic noise, such as pixel artifacts.

A Gaussian filter is used to blur the image, reducing high-frequency noise.

Low-frequency filtering helps reduce the impact of intense temperature changes on the image while preserving the overall thermal pattern.

A median filter is a nonlinear method used to reduce sporadic noise, such as pixel artifacts or "salt and pepper" noise. It works by replacing the value of each pixel with the median of the pixel values in its neighborhood.

The median filtering algorithm is one of the effective methods for reducing noise in thermal images while preserving the clarity of object contours. It starts by selecting a square or rectangular window, for example 3x3 or 5x5 pixels, which is gradually moved across the entire image. For each pixel in the window area, all the intensity values of the pixels that fall into the window are considered. After that, these values are sorted, and the value of the central pixel is replaced by the median of the resulting array.

The main advantage of median filtering is its ability to effectively eliminate pixel artifacts, such as "salt and pepper" noise, without blurring the contours of objects, unlike linear filters. This makes it especially useful for processing thermal portraits, where preserving the shapes and boundaries of objects is critical for subsequent classification and analysis.

At the same time, the method has certain limitations. It requires significant computational resources, especially when working with large images or large windows, which can slow down processing speed. In addition, the median filter is less effective against high-amplitude noise, where individual pixels differ greatly from their neighboring values. Despite this, median filtering remains one of the most reliable methods for cleaning thermal images of small artifacts without losing important information about objectsThe median filter is an effective image processing method for reducing noise and eliminating pixel artifacts without losing clear object contours. The algorithm for its operation involves selecting a square or rectangular window, for example 3x3, which is gradually moved across the entire image. For each pixel, the median value is calculated among all pixels falling into the window: first, all pixel intensity values in the window are sorted, after which the central pixel is replaced by the median of this set.

The main advantages of the median filter are its ability to effectively eliminate pixel artifacts and preserve clear object contours, in contrast to linear filters, which often blur details. This makes it especially useful for processing thermal portraits, where the accuracy of object shapes and boundaries is critical for subsequent classification and analysis.

At the same time, the method has certain limitations. It requires significant computing resources, especially when working with large images or large windows, which can slow down processing. Additionally, median filtering is less effective against high-amplitude noise, where individual pixels differ significantly from their neighboring values. Despite this, median filtering remains a reliable tool for cleaning thermal images of small artifacts without losing important information.

The median of this window will be 60, which will replace the central pixel (100).

The Gaussian filter is a linear filter used to blur images, reducing high-frequency noise. It smooths the image by reducing the sharpness of pixel intensity changes. The Gaussian filter is one of the most common methods of smoothing images, which gives more weight to pixels located closer to the center of the window, and less to pixels on the periphery. The filter algorithm involves using values from the Gaussian filter kernel, which is a two-dimensional matrix. The kernel has the form of a Gaussian function, where the largest value is in the center and gradually decreases to the edges according to a formula that depends on the standard deviation

σ and the distances x and y from the center of the kernel.

During processing, each pixel of the image is replaced by a weighted average value of the pixels in the window, and the weights are determined by the values of the Gaussian filter kernel. This approach allows you to effectively eliminate high-frequency noise, such as random pixel bursts, without significantly reducing the clarity of the contours of objects. The Gaussian filter simultaneously smoothes the image and preserves significant temperature differences, which is especially important when processing thermal portraits for accurate determination of the shapes and thermal characteristics of objects.

Among the disadvantages of the method, it should be noted that if the filter size is too large, important image details can be blurred, which reduces the accuracy of the analysis. Despite this, the Gaussian filter remains an effective tool for removing noise and improving the quality of thermal images, providing a balance between smoothing and preserving key temperature features.

This filter reduces the difference between pixels by replacing the central pixel with a mean value weighted according to a Gaussian function.

Low-pass filtering is a method that allows you to remove high-frequency components of an image while preserving the main low-frequency components that define the overall picture. In the context of thermal images, this helps to reduce the impact of intense temperature changes in the background while preserving the overall thermal picture.

The low-pass filtering algorithm involves using filters that extract the main features of an image, such as a mean or Gaussian kernel. These filters are designed to remove high-frequency details, such as sharp temperature changes at object boundaries or background noise, leaving only the overall thermal patterns of the scene. The most common example of a low-pass filter is the Gaussian filter, which effectively smooths an image by removing high-frequency noise while preserving the main temperature contours.

Among the main advantages of this approach is the ability to effectively smooth

 ISBN 978-3-98924-119-0

the image, eliminating fluctuations caused by noise or short-term temperature changes, while preserving the basic thermal structure of the scene. However, excessive use of low-pass filtering can lead to the loss of important image details, such as small objects or contours, which should be considered when processing thermal portraits for classification and analysis.

Pre-filtering is a crucial step in thermal imaging processing, as it helps eliminate noise, mitigate the effects of sporadic artifacts, and preserve significant temperature changes. A media filter is excellent at dealing with "salt and pepper" noise, a Gaussian filter is good at dealing with high-frequency noise, and low-frequency filtering helps smooth out background temperature changes without losing important data.

 ISBN 978-3-98924-119-0

CHAPTER 3
CONTRAST AND BRIGHTNESS-BASED APPROACHES FOR DETECTING EQUIPMENT

Contrast and brightness are important image characteristics that determine how clearly and distinctly objects stand out against the background. For thermal images, especially in conditions of low temperature contrast between objects and the background, contrast and brightness correction is an important procedure for improving image quality. Thermal imaging cameras often generate images with low contrast, which makes it difficult to perceive details, so various methods are used to normalize and improve contrast.

The histogram equalization method is one of the basic tools for processing images, including thermal imaging. It allows you to analyze the distribution of pixel intensities in an image, i.e., how often different brightness levels occur. This analysis makes it possible to identify patterns in the image, highlight areas with elevated or reduced temperatures, and improve image quality through correction.

A histogram is a graphical representation of the distribution of brightness (or temperature) of pixels in an image. Each bar of the histogram corresponds to a specific temperature range, and the height of the bar shows the number of pixels with that temperature. In other words, a histogram provides a visual representation of how temperatures are distributed throughout the image.

Building a histogram: The brightness (temperature) of each pixel in a thermal image is determined. The brightness values obtained are grouped into intervals, and the number of pixels with that brightness is counted for each interval. The result is a histogram—a graph that shows the distribution of brightness in the image.

The shape of the histogram allows conclusions to be drawn about the contrast of the image, the presence of overexposed or underexposed areas, and the general nature of the temperature distribution. The shape of the histogram can be used to determine how contrasting the image is. A wide histogram indicates high contrast (a large temperature range), while a narrow histogram indicates low contrast. Various anomalies, such as overheated components, can manifest themselves as peaks or dips

 ISBN 978-3-98924-119-0

in the histogram. By analyzing the histogram, you can select the optimal image display parameters to highlight the desired details. By comparing the histograms of different images, you can identify differences in temperature distribution.

The entire temperature range present in the image is divided into several uniform intervals. Next, for each pixel, it is determined to which temperature interval it belongs. Then the number of pixels in each interval is counted. The temperature intervals are plotted on the horizontal axis, and the number of pixels on the vertical axis. For each interval, a rectangle is constructed, the height of which corresponds to the number of pixels.

The histograms shown in the figures allow for a more in-depth analysis of the distribution of temperature values in the thermal image and an assessment of its main properties. If a pronounced peak is clearly visible on the histogram, this indicates that most of the pixels in the frame have a similar temperature corresponding to this maximum value. This situation is typical for homogeneous surfaces or scenes where one type of material or environment prevails.

A wide distribution of values on the histogram indicates that the image has a large temperature range. This may be due to a complex scene containing objects with different thermal characteristics or sharp transitions between warm and cold areas. Such a distribution usually indicates high variability in thermal information, which requires more accurate processing methods for correct interpretation.

An example of a histogram of a thermal image of an uncooled and cooled panel is shown in Fig. 10, 11:

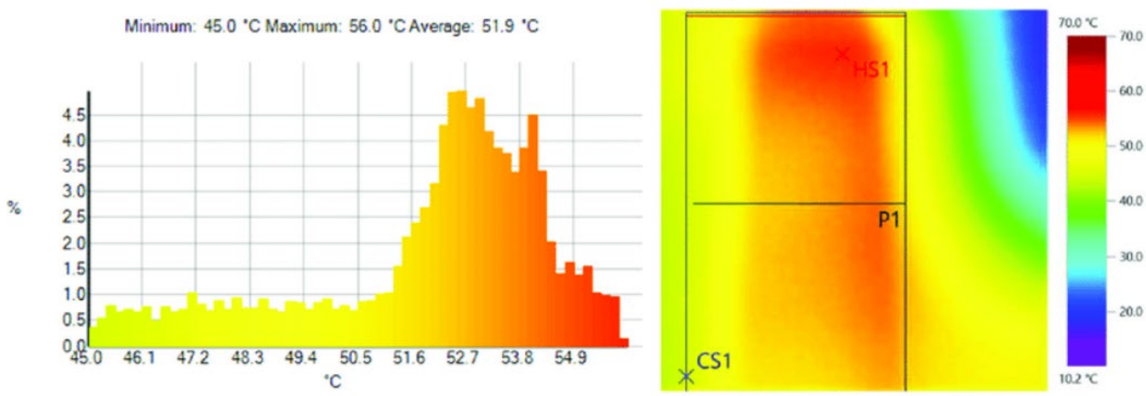

Fig. 10 - Example of a histogram of a thermal image of an uncooled panel

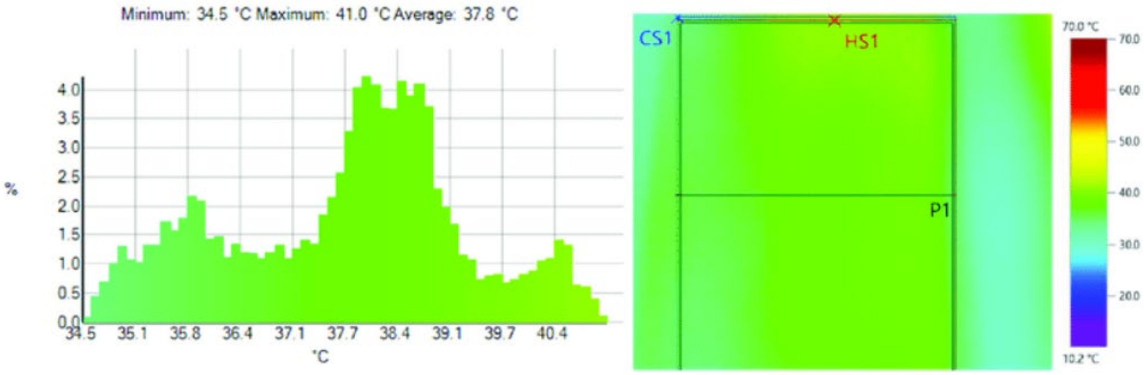

Fig. 11 - Example of a histogram of a thermal image of a cooled panel

In contrast, a narrow distribution indicates a small temperature range in the image. This may mean low contrast, where the difference between warm and cold areas is minimal. In such conditions, the detection of objects or anomalies may be difficult, as temperature differences are practically not recorded by the sensor.

Anomalies in the form of sharp peaks or dips are particularly noteworthy. Such values may indicate the presence of locally overheated or, conversely, abnormally cold areas. Such features are often key in the analysis of thermal portraits, as they can signal dangerous objects, technical malfunctions, heat sources, or camouflage materials that alter the natural thermal background. Thanks to histogram analysis, it is possible to more accurately assess the temperature state of the scene and identify hidden elements that are not always visible in a conventional thermal image.

By analyzing histograms, you can detect defects that may be invisible to the naked eye. Histograms can be used to select the optimal thermal imager settings for a specific task. By comparing the histograms of different objects, you can detect differences in their thermal profiles.

A histogram is an important tool for analyzing thermal images. It allows you to obtain a quantitative assessment of the temperature distribution and identify various anomalies. The use of histograms increases the effectiveness of thermal imaging systems in various fields, from industrial diagnostics to military visualization.

First, a pixel intensity histogram is created for each frame. The histogram shows the number of pixels with specific intensity (brightness) values (from 0 to 255 for 8-bit images), which is an important step for further normalization.

In order to redistribute the intensities evenly, it is necessary to calculate the cumulative distribution of the histogram. The cumulative distribution allows you to see how the intensity of pixels accumulates from the darkest to the brightest pixels.

The cumulative distribution of the histogram (CDH) shows what proportion of pixels in the image have a brightness less than or equal to a certain value. In other words, it is an integral function of the histogram.

The use of the cumulative distribution function (CDF) for uniform brightness redistribution plays an important role in improving the visual properties of an image and ensuring its consistency with other data. First of all, CDF allows linearization, i.e., the conversion of the initial nonlinear brightness distribution into a more uniform and linear one. This simplifies further analysis and allows for better viewing of details that may be hidden or insufficiently clear in an uneven distribution.

Another reason for using CDF is the ability to enhance contrast. By stretching the range of low and high brightness values, the image becomes more distinct: dark areas do not merge with each other, and light areas get clearer boundaries. This visually highlights important elements of the scene, which can be critical for thermal data recognition and analysis tasks.

In addition, the transition to a linear gradation distribution ensures the unification of images with different global brightness levels. This means that after redistribution, they can be correctly compared with each other, regardless of whether one image was taken in conditions of intense heat and the other in a colder environment. Thus, CDF is a universal tool that simultaneously improves the visual quality of an image and ensures its standardization in a series or data set.

The cumulative distribution function (CDF, Fig. 12) shows how pixel intensity values accumulate from the darkest to the brightest pixels. It can be calculated using an image histogram, where each bar of the histogram represents the number of pixels for each intensity. The cumulative distribution is defined as the sum of the histogram values for all intensities, starting with the smallest.

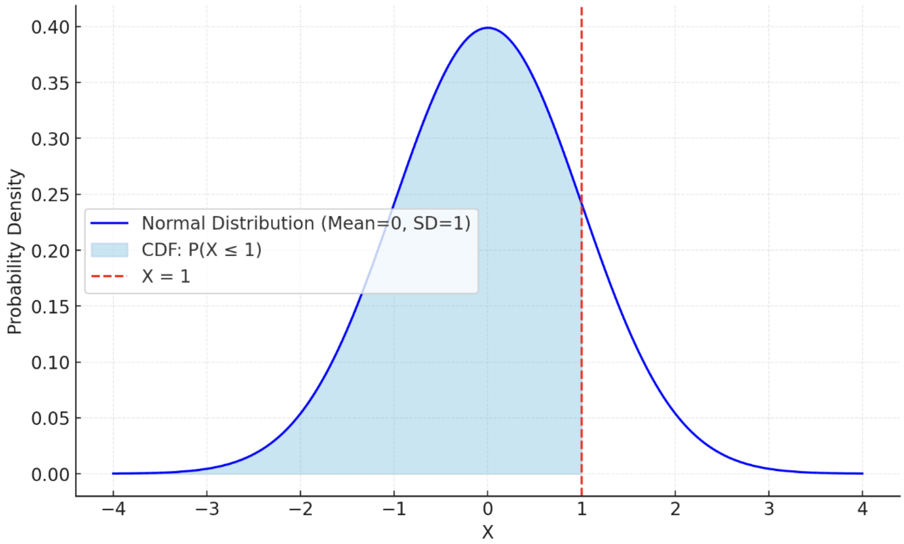

Fig. 12 - Example of a cumulative distribution function

The algorithm for calculating the cumulative distribution function (CDF) starts by constructing a histogram of the image, where $h(i)$ denotes the number of pixels with intensity i. The cumulative distribution function $CDF(i)$ is then calculated, which is the sum of all histogram values up to intensity i. This process allows us to determine the accumulated fraction of pixels for each intensity level, which is an important step in image contrast enhancement methods such as histogram equalization. Using the CDF allows us to evenly distribute pixel intensities across the entire dynamic range, increasing the visibility of details and improving the quality of thermal portraits

CDF calculation algorithm:

1. Construct a histogram for the image, where h(i) is the number of pixels with intensity i.

2. Calculate the cumulative distribution function CDF(i) where CDF(i) is the sum of all histogram values up to intensity i:

$$CDF(i) = \sum_{k=0}^{i} h(k)$$

where:

• h(k) is the number of pixels with intensity k,

• CDF(i) is the sum of the histogram values for intensities from 0 to i.

After calculating CDF, the next step is to normalize this distribution and transform

 ISBN 978-3-98924-119-0

the pixel intensity values. The goal is to redistribute the pixel intensities so that the contrast becomes more uniform and the intensity values cover the entire range from 0 to the maximum possible value (for example, 255 for an 8-bit image).

Pixel conversion algorithm:

1. Calculate the normalized version of the CDF, which will have values in the range from 0 to 1:

$$CDF_{norm}(i) = \frac{CDF_{(i)} - CDF_{min}}{(M \cdot N) - CDF_{min}}$$

where:

• CDFmin is the minimum value of the cumulative distribution (for example, for low-contrast images, this may be 0),

• M·N is the total number of pixels in the image (for an image of size M×N).

2. We convert the CDF values into new intensities using the normalized function:

$$s(i) = round\ (CDF_{norm}(i) \cdot 255)$$

where:

• s(i) — the new intensity value for the pixel with intensity i.

After each pixel in the image has been converted to a new intensity value, the image will have improved contrast, and details that were previously difficult to distinguish will become more pronounced. The cumulative distribution of the histogram is a powerful tool for correcting contrast and improving the quality of thermal images. However, its use requires caution to avoid losing important information. Here is an example of how the histogram of an image changes before and after histogram equalization (Fig. 13).

Before histogram equalization: the image histogram may have a cluster of pixels in one range of intensity values, resulting in low contrast.

After histogram equalization: the histogram becomes more uniform, and the pixel intensity values are evenly distributed across the entire range from 0 to 255, which improves contrast.

You can also use Adaptive Histogram Equalization (AHE).

 ISBN 978-3-98924-119-0

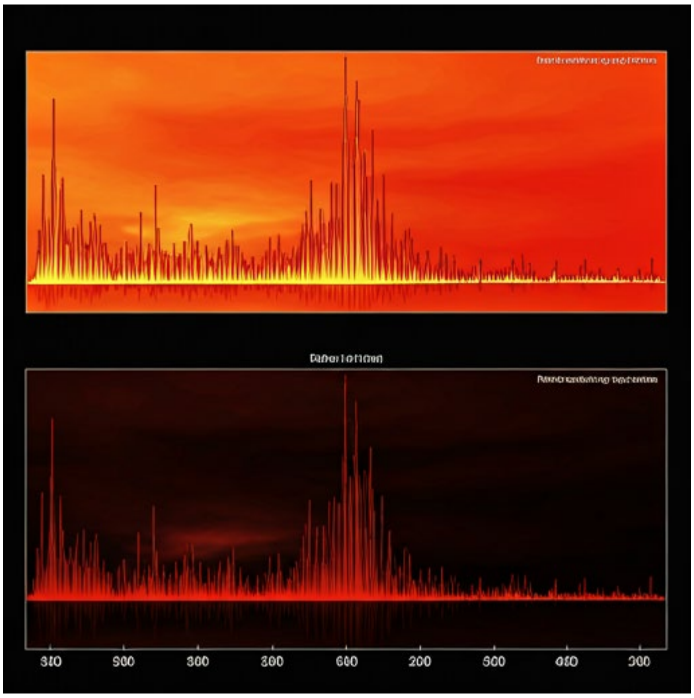

Fig. 13 - Modified histogram

Adaptive histogram equalization is an advanced method used to improve contrast in specific areas of an image. Instead of distributing contrast evenly across the entire image, AHE allows you to increase contrast only in local areas, which is especially useful when the image has different brightness or texture levels in different parts of the image.

The algorithm for adaptive histogram equalization begins by dividing the image into small sub-areas, or local blocks, for each of which a histogram is calculated. Histogram equalization is then performed independently for each block, adjusting the contrast based solely on the local distribution of pixel intensities within that area. Following this, the pixel values within each block are updated according to the results of the local histogram equalization. Finally, after processing all blocks, the local areas are merged to reconstruct a single global image with enhanced contrast, providing improved visibility of details across regions with varying lighting or thermal intensity.

If you imagine conventional histogram equalization as uniformly stretching a rubber band across its entire length, adaptive equalization is a more localized approach. Instead of changing the brightness of all pixels in the image equally, this method divides the image into smaller blocks and aligns the histogram of each block separately. This preserves local contrasts and details that could be lost with global alignment.

Thermal images often have uneven brightness, especially when capturing objects with different temperatures. Traditional histogram equalization can result in Adaptive equalization is an effective method for enhancing thermal images, particularly in cases where overexposed or underexposed areas obscure important details.

This technique allows for the preservation of fine details in regions with both high and low contrast, making it possible, for example, to highlight small defects against larger objects. Additionally, adaptive equalization compensates for uneven lighting or temperature variations across different parts of the image, ensuring that all regions are represented clearly. By improving the overall contrast and balance, this method enhances visual interpretation, making thermal images more comprehensible and informative for human observers.

Thermal images often contain areas with different temperatures, resulting in significant variations in brightness. This can complicate visual analysis and defect detection. Adaptive equalization allows you to enhance the contrast in each local area of the image, making details more visible.

Before alignment: The image may contain areas with very high temperatures and areas with low temperatures. Due to such a wide temperature range, details may be lost in shadowed or overexposed areas.

After alignment: Adaptive alignment allows you to enhance the contrast in each local area of the image. This means that both hot and cold areas become more visible.

Examples of thermal images before and after adaptive histogram equalization presented on Fig. 14.

Adaptive histogram equalization has a number of important advantages, thanks to which this method is widely used in image processing tasks. Its main feature is the ability to preserve local details: instead of global brightness redistribution, the algorithm works separately with small areas of the image, allowing you to clearly highlight fine contours, textures, and local temperature features. Another important advantage is its adaptation to uneven lighting or a non-uniform thermal background. The method effectively processes scenes where one part of the image is overexposed and the other is in partial shade, ensuring balanced detail reproduction.

 ISBN 978-3-98924-119-0

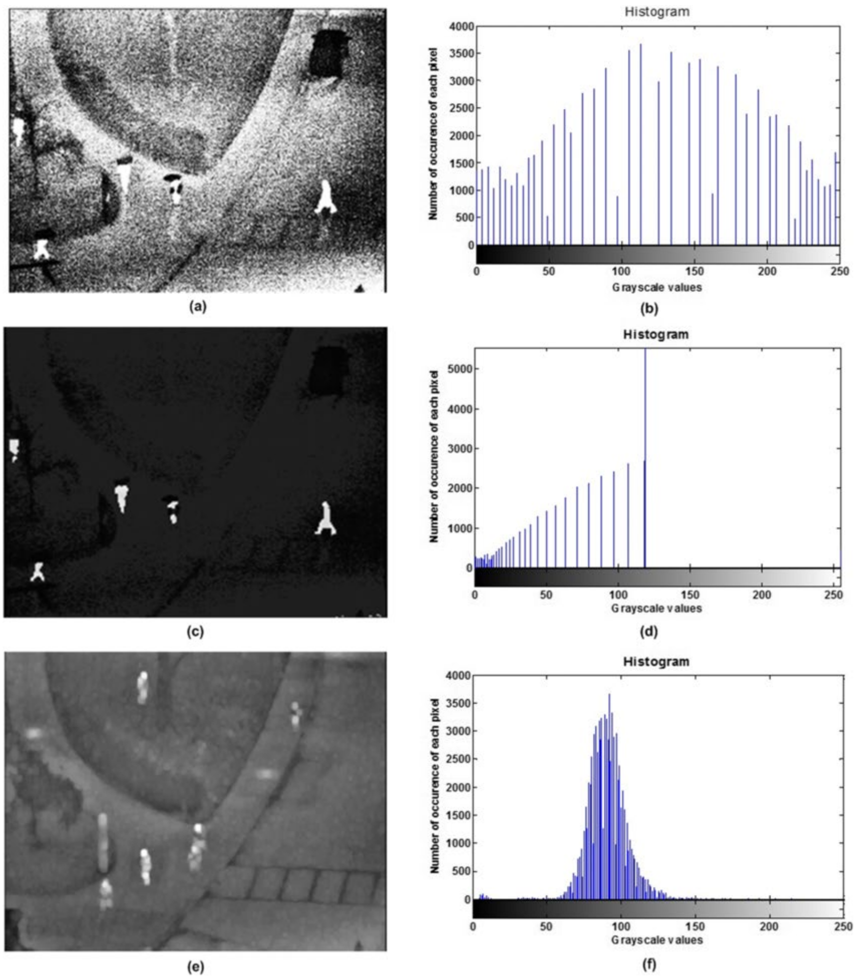

Fig. 14 - Examples of thermal images before and after adaptive histogram equalization

Description: The difference in temperature between different areas of the wall has become more apparent, allowing heat loss locations to be identified.

By enhancing local contrast, adaptive histogram equalization significantly improves the overall visual perception of the scene. Objects become clearer, differences between areas are easier to recognize, which is especially important for analyzing thermal images and other data with low initial contrast. The method is universal and is used in many fields, from medicine and unmanned systems to military intelligence and technical diagnostics.

At the same time, there are certain drawbacks. Since processing is performed in blocks, artifacts can sometimes occur at the boundaries of these blocks—noticeable transitions or stripes that distort the naturalness of the image. In addition, the method requires careful adjustment of parameters such as block size and contrast coefficients.

An unsuccessful choice of parameters can either not enhance the details sufficiently or, conversely, create an overly contrasting and unnatural image.

Adaptive histogram equalization is a powerful tool for improving the quality of thermal images. It allows you to extract maximum information from the data obtained and make it more understandable for analysis.

You can use the contrast-limited adaptive histogram equalization (CLAHE) method. CLAHE is an improved version of AHE that limits contrast enhancement for each region, avoiding excessive noise amplification. In CLAHE, a histogram is calculated for each region, but there is a limit on the maximum number of pixels that can be transferred to a certain range of intensity values.

CLAHE (Contrast Limited Adaptive Histogram Equalization) algorithm visualization on Fig. 15:

1. Image segmentation into blocks. The input image is divided into uniform blocks (tiles) of fixed size. This allows local contrast enhancement instead of global processing.

2. Calculation of the histogram for each block. For each tile, a brightness distribution histogram is formed, which reflects how often certain intensity levels occur in this local region.

3. Clipping the histogram. Histogram bin values that exceed the set threshold (clip limit) are reduced to the maximum allowed level. The excess ("clipped excess") is distributed evenly among all bins, which avoids excessive contrast enhancement and noise.

4. Local histogram equalization. For each block, a cumulative distribution function (CDF) is calculated based on the adjusted histogram. The pixel values in the block are transformed according to the CDF, which ensures a more uniform brightness distribution and improved local contrast.

5. Bilinear interpolation between blocks. To avoid sharp transitions between the boundaries of neighboring blocks, bilinear interpolation is used for pixels in transition areas. This ensures a smooth merge of regions with different contrast levels and prevents the appearance of artifacts.

 ISBN 978-3-98924-119-0

6. Formation of the final image. After processing all blocks and applying interpolation, the results are combined into a final image that has enhanced local contrast and preserves the naturalness and smoothness of tonal transitions.

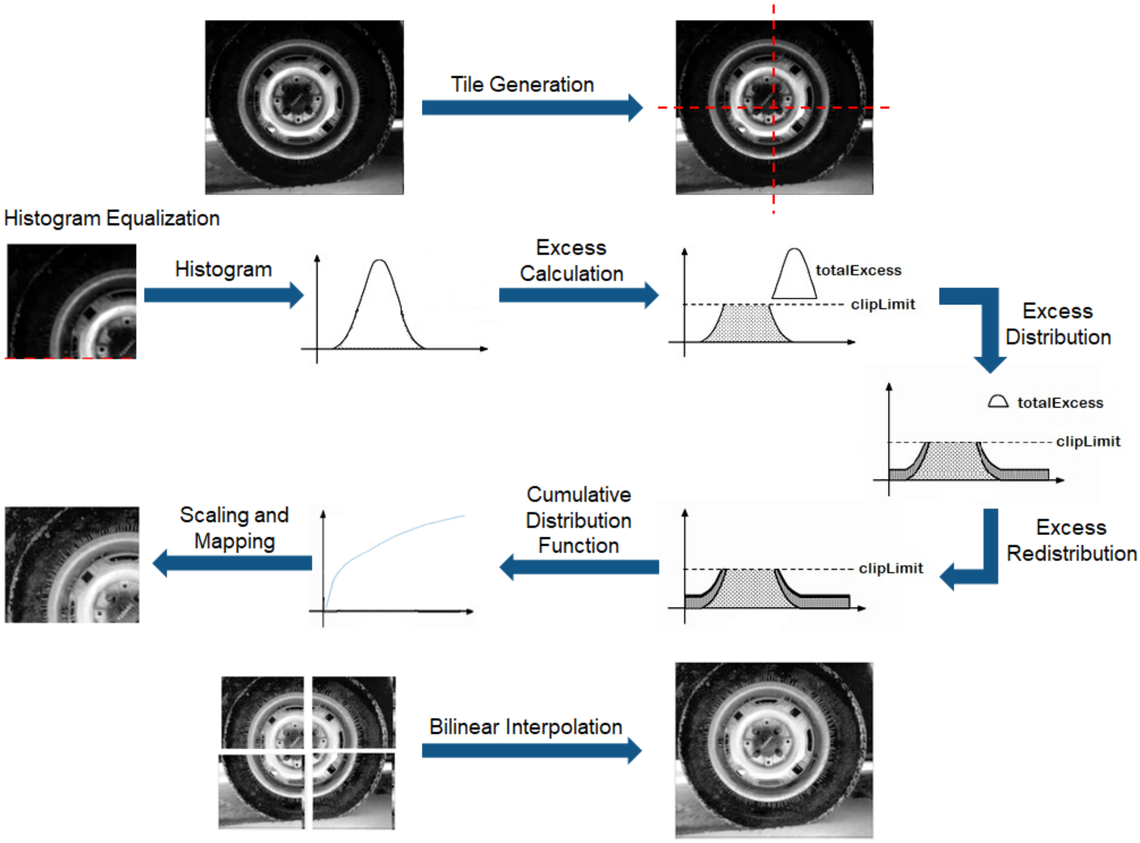

Fig. 15 - Contrast Limited Adaptive Histogram Equalization algorithm visualization

The CLAHE method is an excellent choice for improving the contrast of thermal images of military equipment. CLAHE allows you to enhance the contrast in each local area of the image, which is especially useful for heterogeneous thermal signatures of military equipment (e.g., hot engines, cold metal surfaces). The method prevents oversaturation of bright areas, allowing details to be preserved in high-contrast areas. CLAHE can help reduce the impact of noise that is often present in thermal images. Small details such as weapons, antennas, or people near equipment become more visible.

Contrast-limited adaptive histogram equalization (CLAHE) has a number of advantages over the classical histogram equalization method. One of the main advantages is the ability to preserve local image details: CLAHE increases the contrast in each individual area without negatively affecting other parts of the image. This is especially important when processing thermal portraits, where small details can contain critical information about objects or anomalies.

Another advantage is that it is less prone to artifacts, such as sharp brightness drops at block boundaries, which often occur when using classical histogram equalization. CLAHE allows for smoother contrast changes, providing a more natural look to the processed image. In addition, this method adapts better to uneven lighting and heterogeneity of thermal signatures of objects, which makes it especially useful for analyzing complex thermal scenes, where the temperature of different areas can vary significantly. Thanks to these properties, CLAHE is an effective tool for improving the quality of thermal images and the accuracy of their further processing.

The CLAHE method is a powerful tool for improving the quality of thermal images of military equipment. It allows you to obtain more detailed and contrasting images, which facilitates their analysis and interpretation.

Contrast and brightness correction are important steps in thermal imaging processing because they improve the visibility of objects against the background. The histogram equalization method is useful for general contrast enhancement, but it can lead to loss of detail in the extreme areas of the image. Adaptive histogram equalization and CLAHE allow for more precise contrast adjustment for local areas, which is especially useful for complex thermal maps with significant temperature variations.

Recognizing and highlighting objects in thermal images is an important step in data processing. When analyzing thermal maps, it is important to correctly select objects such as people, vehicles, equipment, etc., so that they can be tracked, classified, or used for further tasks such as forecasting or analyzing thermal anomalies. Various segmentation and contour selection methods are used for this purpose.

Image segmentation is the process of dividing an image into separate areas or segments that share common properties such as color, intensity, or texture. For thermal images, segmentation methods based on temperature thresholds are most often used , since thermal objects usually have a higher or lower temperature than the background.

Threshold segmentation is a basic method that involves determining whether each pixel in an image belongs to an object or the background, depending on its temperature value. If the temperature of a pixel exceeds a certain threshold, that pixel is considered part of the object. If the pixel temperature is below the threshold, the pixel belongs to the background.

The threshold segmentation algorithm is one of the simplest and most common methods for processing thermal images. At the first stage, a temperature threshold ($T_{threshol)d}$ is determined, which can be set manually or calculated using statistical methods, for example, from the analysis of the temperature histogram in the image. After setting the threshold, each pixel is processed individually: if the pixel intensity $I(x, y)$ is equal to or exceeds the threshold, it belongs to the object, and if the intensity is less than the threshold, the pixel belongs to the background.

The main advantages of this approach are the simplicity of implementation and high processing speed, which makes it attractive for real-time tasks or processing large arrays of thermal images. At the same time, the algorithm has certain limitations. Accurately setting the threshold is critical, since an incorrect value can lead to incorrect selection of objects or inclusion of parts of the background in the segmented area. In addition, the method is not effective in cases where the temperature of the object and the background are very close, as this makes it difficult to separate them and can reduce the accuracy of classification.

Adaptive segmentation is a more flexible method than threshold segmentation. In this case, the temperature threshold is determined not for the entire image, but for each local area separately. This allows you to take into account different lighting conditions or different temperature fluctuations in different parts of the image. Local thresholds can vary depending on the temperature dynamics in the frame, which makes the method more resistant to changes in shooting conditions.

Adaptive segmentation algorithm is an effective method for processing thermal images, as it allows to take into account local changes in temperature intensity in different parts of the scene. At the first stage, the image is divided into small areas, for example, 8x8 or 16x16 pixels in size. For each of these areas, a local threshold T_{local} is calculated, which can be defined as the average temperature value, median or other value calculated using adaptive methods.

After determining the local thresholds, the pixels in each area are segmented. Pixels whose temperature exceeds the local threshold are attributed to the object, while pixels with a temperature below the threshold are attributed to the background. This approach allows for more accurate object separation, especially in cases where the temperature of different parts of the scene is significantly different or there are local anomalies.

The advantage of adaptive segmentation is the ability to take into account local differences in temperature intensity, which increases the algorithm's resistance to changes in lighting, thermal noise or anomalies in thermal images. However, this method has a higher computational complexity compared to standard threshold segmentation. In addition, the choice of region size is critical: regions that are too large can lead to smoothing of small objects, and regions that are too small can lead to increased noise. Therefore, adaptive segmentation requires a balance between object detection accuracy and computational efficiency.

Once objects have been segmented, the next step is to clearly highlight them in the image. Edge detection allows you to obtain clear boundaries of objects, which is important for their further classification and analysis. One of the most common methods for edge detection is the Canny Edge Detection method, which is a classic approach to determining edges in images.

The Kenny edge detection algorithm is one of the most widely used approaches to extracting features from images due to its high accuracy and noise immunity. In the first step, the image is smoothed using a Gaussian filter, which reduces noise and removes small details that can create false edges. This is a preparatory step necessary to increase the reliability of subsequent analysis.

In the second step, the gradient is calculated for each pixel to determine the direction of the sharpest intensity change. Sobel or Prewitt operators are commonly used for this purpose, as they allow computing the horizontal and vertical components of the gradient. This step makes it possible to identify areas of rapid intensity variation, which typically correspond to object edges.

Next, two thresholds—high and low—are applied. Pixels whose gradient magnitude exceeds the high threshold are classified as edge pixels, while those below the low threshold are discarded. Pixels with gradient values between the thresholds are included in the contour only if they are connected to strong-edge pixels (those above the high threshold). This hysteresis-based approach improves the stability of contour detection and reduces the impact of noise.

Finally, fine smoothing may be applied to the extracted contours to eliminate minor artifacts caused by small local fluctuations in intensity. The Canny method (referred to here) is valued for its ability to sharply delineate object boundaries and maintain robustness when processing noisy or complex thermal images. At the same time, its limitations include difficulty detecting weak or low-contrast contours and sensitivity to the selected threshold parameters, which must be carefully tuned for optimal performance.

The impact of contrast enhancement on the quality of edge extraction, especially when applying the Contrast Limited Adaptive Histogram technique, is illustrated in Figure 16. As shown, the improved local contrast achieved through CLAHE enables the Canny detector to identify edges more accurately and consistently, enhancing the stability of contour extraction in thermal imagery and revealing object boundaries that would otherwise remain indistinct.

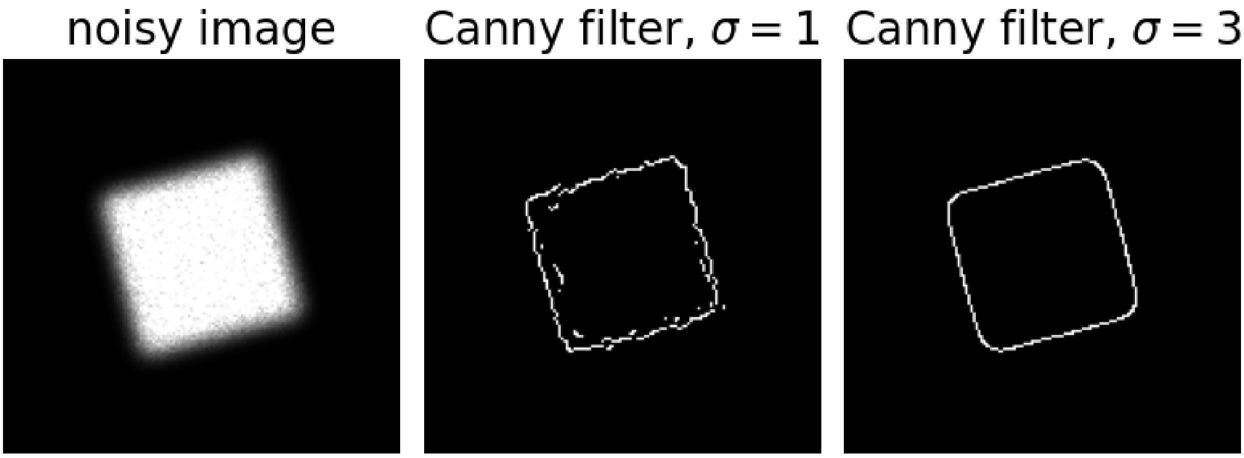

Fig. 16 - Canny Edge Detection Method Example

Threshold segmentation is a basic method, but for more complex situations, adaptive segmentation, which takes into account local temperature variations, is more

ISBN 978-3-98924-119-0

effective. Contour detection, particularly using the Kenny method, enables the acquisition of clear object boundaries, which enhances their subsequent processing and analysis.

Image geometry correction is a crucial step in thermal imaging data processing, as thermal imaging cameras can be installed at varying angles or in conditions where the image is not uniform. This can lead to distortion of object proportions or incorrect display of temperature values. To achieve accurate results and maintain correct proportions and temperature measurements, image geometry correction methods must be applied, in particular perspective correction and calibration.

Perspective correction is necessary when the thermal imager is positioned at an angle to the observed object. In such cases, the image of the object may be distorted: straight lines may appear curved or uneven, and the proportions of objects may change. For example, if the thermal imager is not directed directly at the object, but at an angle, the front parts of the object will appear larger than the rear parts, which distorts the correct proportions.

The perspective correction algorithm is a crucial tool in thermal image processing, as it enables the restoration of the correct geometry of objects distorted by the camera's viewing angle. The first step of this algorithm is to determine the corresponding points in the image. These are pixels or characteristic elements that have known geometric properties or are easily identified, for example, object corners, lines or borders. The selection of such points is critical, since it is on their basis that distortion correction will be carried out.

After determining the corresponding points, a perspective transformation is applied, which changes the coordinates of the image pixels so that the specified points are in the correct positions. This transformation is described by a matrix that defines how the coordinates of each pixel of the original image are transformed to achieve the correct appearance. As a result of processing, the image acquires the proper proportions and geometry of objects, which significantly facilitates further classification and analysis of thermal portraits.

The basic formula for perspective transformation is as follows:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Where:

• x, y, z — coordinates of the pixel in the original image,

• x', y', z' — coordinates of the pixel in the corrected image, and matrix

• H is the homography matrix, which defines the transformation for perspective correction.

Applying a homographic matrix to an image is an important step in perspective correction, especially in cases where objects in thermal portraits are distorted due to the shooting angle. After calculating the homographic matrix, it is applied to each pixel of the image, which allows you to obtain a corrected picture with the correct proportions and geometry of objects. This is especially useful for analysis and classification, as it allows you to accurately estimate the size, shape, and relative location of targets.

Among the main advantages of this approach is the correction of distortions caused by the viewing angle, as well as the restoration of reliable proportions of objects in the image. This increases the accuracy of further analysis and makes the classification results more reliable. At the same time, the use of a homographic matrix has certain limitations. For high-quality correction, it is necessary to accurately determine the corresponding points in the image, which can be difficult in the case of low-contrast or fuzzy thermal portraits. In addition, the transformation sometimes leads to the loss of some information at the edges of the image, which should be taken into account when processing large or important areas of the scene.

Calibrating a thermal imager is the process of adjusting temperature measurements to account for factors that can affect the accuracy of the results, such as ambient temperature, object type, or the settings of the thermal imager itself.

Thermal imagers work by detecting infrared radiation emitted from objects. However, different objects can emit infrared waves differently, depending on their material, condition, and other factors. Calibration allows these factors to be taken into account and ensures the accuracy of temperature measurements.

The ambient temperature can affect the measurement results, as the thermal imager itself can also heat up, or its sensor may have an offset that depends on the temperature in its environment. Calibration allows these effects to be reduced or eliminated.

Thermal imagers may have different sensitivities to materials, as different materials have different emission coefficients. Calibration helps to take these characteristics into account and adapt the thermal imager to different types of surfaces.

Special calibration standard objects, such as black bodies with a known temperature or special test plates with specific properties, can be used for calibration. They are used to compare the results of the thermal imager and ensure the accuracy of measurements.

After identifying errors and deviations, correction factors are applied to change the measured temperature to match the actual value. This may include corrections based on ambient temperature, material, or viewing angle.

After calibration, the thermal imager is periodically checked and adjusted to ensure consistent measurement accuracy, as factors such as ambient temperature can change over time.

Calibration of thermal imaging systems is an important step in ensuring the accuracy of measuring the temperature of objects in different conditions. One of the main advantages of calibration is the ability to obtain reliable temperature readings, regardless of external factors or operating conditions. It allows you to take into account various operating parameters of the thermal imager, as well as the material properties of the objects being observed, ensuring a more accurate interpretation of thermal portraits and increasing the reliability of the analysis results.

However, calibration also has certain limitations. The procedure requires specialized equipment and time, which can make it difficult to carry out in field conditions or in operational mode. In addition, re-calibration may be required for each new situation or new object, since changing conditions or properties of the observed object can affect the accuracy of measurements. Thus, although calibration significantly increases the accuracy of thermographic data, its application requires

 ISBN 978-3-98924-119-0

careful planning and resources.

Perspective correction helps restore the correct proportions of objects, even if the thermal imager is at an angle to the object, and calibration ensures the accuracy of temperature measurements, taking into account various factors such as ambient temperature and material properties of objects.

Improving image quality is an important aspect of video processing, especially for thermal images, which often suffer from low resolution or insufficient detail. To address these issues, interpolation and super-resolution techniques are used to enhance image quality while preserving important details.

Spatial interpolation is used to increase the number of pixels in an image, thereby improving its resolution. This process allows for greater detail by adding new pixels between existing ones, creating an image with higher resolution.

Interpolation methods: linear interpolation, bicubic interpolation, nearest neighbor interpolation.

Linear interpolation uses the average value between neighboring pixels to calculate a new pixel. This is a simple method that gives more or less uniform results, but may not be accurate enough for images with great detail.

Pixel interpolation is a crucial technique in image processing, particularly for enhancing resolution and smoothing pixel transitions. For two pixels, p1 and p2, the calculation can be performed in different ways, depending on the chosen interpolation method. One of them is bicubic interpolation, which uses four pixels instead of the two used in bilinear interpolation. This enables smoother transitions and better preservation of image details, particularly when the image features highly detailed objects. Although bicubic interpolation is more computationally demanding, it provides high-quality processing and is often used to enhance thermal portraits, where preserving minor thermal differences is essential.

Another, more straightforward approach is nearest neighbor interpolation, in which a new pixel copies the value of the closest available pixel. This method is faster, but can result in coarse pixel transitions and loss of detail, which limits its effectiveness for high-precision thermal image analysis tasks.

The main advantages of interpolation methods are the ability to obtain smoother and more detailed images, with bicubic interpolation being particularly effective when working with highly detailed frames. Disadvantages include increased processing time, especially for large images, and the fact that if artifacts are present in the original photos, interpolation may not yield ideal results, sometimes exacerbating existing distortions.

Super-resolution is a method that increases the resolution of an image or video by analyzing multiple frames or images. This method allows you to recover details that are missing due to the limited resolution of the original image. Super-resolution often utilizes multiple images to gather more information and create a higher-resolution image with greater detail.

Multi-frame super-resolution: In this approach, multiple frames of the same object or scene are analyzed to gather additional information about details that may be lost in a single image. By comparing these frames, algorithms can extract missing information and create new pixels to increase the resolution.

Frame registration (or alignment) algorithms are used for precise image positioning, enabling high accuracy when combining frames.

Deep learning-based super-resolution: Recently, super-resolution methods that utilize neural networks to analyze images have gained popularity. Neural networks can learn from large image datasets to restore missing details and increase image resolution. Algorithms such as SRGAN (Super-Resolution Generative Adversarial Network) use a generative adversarial network approach to generate high-quality images.

Multi-frame super resolution: This method utilizes multiple frames of the same object or scene, taken from different angles or under varying lighting conditions. Algorithms collect information from each frame to improve the overall image quality by increasing its resolution.

Super-resolution thermal image processing offers several advantages that make it a valuable tool for enhancing the quality of thermal portraits. Thanks to super-resolution processing methods, it is possible to obtain high-quality images even with limited sensor resolution. Deep learning-based algorithms are particularly effective, as they can restore details that are difficult or impossible to recover using traditional image processing methods. This enables more accurate object detection, analysis of their shape and structure, and enhances the reliability of subsequent classification.

At the same time, super-resolution processing has certain limitations. It requires significant computing resources, particularly when working with large datasets or utilizing deep neural networks. In addition, the quality of the final result largely depends on the source frames: if the image is severely degraded or contains artifacts, the restored details may be inaccurate, which limits the effectiveness of the method under challenging conditions. Thus, although super-resolution processing significantly improves the quality of thermal images, its application requires a balanced approach that considers both computational resources and the quality of the source data.

Interpolation and super-resolution are potent tools for enhancing the resolution and detail of images, particularly in low-quality or low-resolution conditions. Spatial interpolation helps increase the number of pixels and improve detail, while super-resolution allows for increased resolution through the analysis of multiple frames or the use of neural networks. These methods can significantly improve image quality and make images more accurate for further study and recognition.

Integrating data from different sensors is a crucial step in enhancing the accuracy of information processed by a single sensor, such as a thermal imager. Since each type of sensor has its limitations, combining data from different sources can compensate for these shortcomings and provide a more accurate and comprehensive picture of the situation. This is especially important in situations where it is necessary to accurately detect, identify, or track objects, such as in military, security, or search-and-rescue scenarios.

In many cases, a combination of different sensor types is used to obtain more accurate data about the environment or the objects being observed. For example, objects can be detected or identified using a thermal imager. Still, for more precise identification and classification, it is advisable to combine this data with other sources, such as optical cameras, radar systems, lidar, or acoustic sensors.

CHAPTER 4
FEATURES OF COMBINING OPTICAL AND THERMAL APPROACHES IN EQUIPMENT DETECTION AND CLASSIFICATION TASKS

Optical cameras capture images within the typical visible spectrum, enabling a clear picture of objects in real-time. They help to better identify objects by providing more context and detail that may be lacking in thermal images, especially in low-temperature contrast conditions. For example, in low-light conditions or poor visibility, thermal imagers may be less effective, whereas optical cameras can provide high-quality images under certain lighting conditions.

Radars use electromagnetic waves to detect objects, even in challenging weather conditions (such as rain, fog, or snow). They enable the detection of objects at long distances and are highly resistant to atmospheric influences. Radars can be particularly useful for detecting large or difficult-to-see objects such as aircraft, cars, or ships, as their ability to operate in all conditions provides essential information that cannot be obtained with thermal imaging cameras or optical cameras.

Lidar (Light Detection and Ranging) uses laser pulses to determine the distance to objects and create three-dimensional maps of the surrounding environment. This sensor provides accurate data on the shape, size, and location of objects, which is essential for building 3D models of the territory or for studying complex objects such as buildings or landscapes. Lidar can complement thermal imaging data, especially in cases where accurate distance measurement or object dimensions are required.

Combining data from different sensors, also known as sensor fusion, is a crucial step in achieving more accurate and reliable results. Sensor fusion involves processing data obtained from various sources to achieve more accurate results.

The correlation method is one of the simplest ways to integrate data. It involves correlating data from different sensors based on similar features or parameters. For example, thermal images and images from an optical camera can be processed to find common points based on similar structures or shapes.

Kalman filtering is a powerful mathematical method that enables the combination of data from different sources, taking into account not only accurate measurements but

 ISBN 978-3-98924-119-0

also the uncertainties associated with each source. The Kalman filter enables the estimation of an object's state by combining data from sensors in a manner that minimizes the impact of errors and noise in each source. This is especially useful when using sensors with different accuracies and sensitivities.

For more complex scenarios, machine learning algorithms are often used to analyze and integrate data from multiple sensors automatically. For example, neural networks can be used to study patterns and properties of objects that different sensors can detect and to make decisions based on multi-sensor data.

Combining different sources of information reduces errors that can occur when using only one type of sensor. For example, objects that are difficult to see in a thermal image can be clearly identified using data from an optical camera or radar.

The integration of various sensors enables operation in diverse conditions. If one sensor (e.g., an optical camera) cannot operate in darkness or poor visibility, other sensors, such as radar or thermal imaging, can fill the gap.

In conditions where a single sensor may be prone to noise or emissions, sensor fusion enables a more stable and reliable image or data by combining multiple sources of information.

Processing and synchronizing data from multiple sensors can be computationally complex and resource-intensive, which can slow down the processing. Proper integration requires the use of complex algorithms that can accurately process and combine data from different sources, necessitating high precision and in-depth knowledge in the fields of signal processing and machine learning.

Integrating data from multiple sensors is a powerful tool for enhancing object detection accuracy and situational analysis. Combining thermal imaging data with other sensors, such as optical cameras, radar, or lidar, provides a more comprehensive and accurate picture, thereby reducing the likelihood of errors and enhancing the reliability of the results. Sensor fusion not only enhances detection efficiency but also provides improved adaptation to various operating conditions.

Developing an interface for displaying results is an essential step in creating an effective system for processing data obtained from thermal imaging cameras. The

 ISBN 978-3-98924-119-0

primary purpose of such an interface is to provide the user with clear, convenient, and intuitive information. This may include thermal map visualization, object tracking, and additional features that allow the user to quickly and effectively respond to situations that require attention.

Thermal maps are the primary method for visualizing temperature data obtained from thermal imaging cameras. They allow you to clearly see temperature fields and highlight areas with elevated or reduced temperatures. The interface for displaying such maps should be simple, yet contain all the necessary tools for accurate analysis.

Thermal maps use different colors to represent temperature values, where red or yellow usually represent high temperatures and blue represents low temperatures. The palette must be clearly visible and understandable to the user. Determining the correct contrast between colors enables the highlighting of essential temperature anomalies.

Each thermal map should indicate which temperatures correspond to specific colors. For example, the scale ranges from 0°C to 100°C, where each color on the map corresponds to a particular temperature within this range. For accuracy, the interface should have a temperature scale displayed next to the map.

For user convenience, areas with elevated temperatures or anomalies should be automatically highlighted or illuminated. This may include the automatic marking of objects such as equipment, people, or other potentially essential objects that stand out against the background due to temperature changes.

Heat maps can be dynamic, meaning they are updated in real-time with new data. The interface should be designed to allow users to observe changes in temperature fields without delay, with the ability to zoom in and analyze specific points.

The user should be able to interact with the thermal map, for example, by changing the scale, moving the image, zooming in or out on specific areas of the image. The interface may include functions for saving or exporting maps in various formats (e.g., .png, .jpg, .tiff).

Tracking objects in thermal images is crucial for monitoring and predicting their movement. To achieve this, it is necessary to develop an interface that enables tracking the movement of objects across frames and provides real-time predictions of their

 ISBN 978-3-98924-119-0

movement.

The first step in tracking is to identify objects that require attention in each frame. This may include automatic object detection based on temperature anomalies or contour analysis. For more complex situations, computer vision or machine learning methods can be used for more accurate identification.

Tracking algorithms, such as the Kalman filter or the SORT (Simple Online and Real-time Tracking) algorithm, enable the tracking of an object's movement from frame to frame, even when the object temporarily disappears or moves quickly. This is important for accurately identifying its trajectory and predicting its next position.

The user should be able to see not only the current position of the object, but also its previous trajectories on a heat map. This can be done using different colored lines or markers that show the movement of the object over a specific period of time.

Additionally, the system can integrate prediction algorithms that forecast the future location of an object based on its current movement. This can be useful for monitoring decisions or for taking operational measures in situations where a quick response is essential.

The user should be able to configure tracking parameters, including selecting objects for tracking, changing algorithms, and filtering objects by specific criteria (e.g., temperature values or sizes). Also, the interface should allow you to turn on or off tracking of individual objects, as well as provide statistics on their movement (speed, direction, predicted location).

To make the system user-friendly, the interface should include interactive features for detailed analysis of detected objects and their movement:

Zooming and panning allow the user to enlarge the image for a more accurate analysis of specific areas.

The ability to select a specific time interval for displaying data, which makes it possible to analyze changes over a given period.

The interface can include functions for automatically highlighting abnormal temperature changes, enabling the user to draw attention to important details quickly.

Visualization of heat maps helps users clearly see temperature fields, and object

ISBN 978-3-98924-119-0

tracking enables them to track the movement of objects in real-time, which is helpful for monitoring and forecasting. The interface should be intuitive and include flexible configuration features, allowing the system to be adapted to different conditions and user needs.

Optimizing algorithms for real-time operation is a critical aspect of developing systems that must process large amounts of data, such as thermal imaging cameras or machine learning-based systems. In real time, the algorithm must process data instantly to ensure accurate results without delay. This requires the development of effective methods that optimize both the speed and accuracy of data processing. One way to achieve these goals is to use specialized hardware and machine learning algorithms.

One of the most effective ways to optimize algorithms for real-time performance is to utilize hardware that significantly accelerates data processing. Essential components of hardware acceleration are graphics processing units (GPUs) and distributed computing systems, which can process large amounts of data much faster than traditional central processing units (CPUs).

Graphics processing units (GPUs) are specialized chips initially designed to process graphics, but have since proven highly effective for parallel computing. Thanks to their large number of computing cores, GPUs are capable of processing thousands of data streams simultaneously, making them ideal for tasks that require high processing speeds, such as real-time video or image processing.

GPUs can perform thousands of operations simultaneously, which speeds up the processing of large amounts of data. This is important for image processing tasks, where each pixel may require separate processing.

Thanks to their specialized architecture, GPUs can be significantly faster than traditional CPUs in tasks that require a large number of calculations to be performed simultaneously.

GPUs are widely used for training neural networks due to their ability to process large data sets and perform the mathematical operations required for model training.

For a system that processes video streams from thermal imaging cameras, a GPU can be utilized to accelerate operations such as image filtering, segmentation, object

 ISBN 978-3-98924-119-0

detection, and complex computations on large datasets in real-time. A GPU allows each video frame to be processed without significant delays, which is necessary for real-time applications.

Another way to optimize algorithms for real-time is to use distributed computing systems. Distributed systems utilize multiple computers or servers to process data simultaneously, thereby significantly increasing computing power.

With distributed computing, you can scale the system according to the amount of data. This means that you can add new computing power as needed to handle large amounts of data.

Distributed systems can be more reliable because if one component fails, the system continues to operate on other computers.

By distributing computing tasks across multiple machines, processing time can be significantly reduced.

In systems that analyze video in real-time, distributed computing systems can process different video frames on other machines, thereby reducing the time required to process a large video stream.

Machine learning and neural networks are powerful tools for detecting and classifying objects in images or videos. The use of machine learning algorithms enables the automation of object analysis and classification, significantly improving the accuracy and speed of real-time data processing.

Neural networks, particularly convolutional neural networks (CNNs), are widely used for image processing because they are capable of detecting complex patterns and structures in data. For real-time video processing, neural networks can be utilized to automatically detect objects, classify them, and predict their movement trajectories.

Neural networks can learn independently from large data sets, allowing them to improve the accuracy of object classification and detection over time.

Neural networks can be configured to solve a variety of tasks, including object detection, image segmentation, and prediction.

Networks can work with large datasets and process information in real-time when using appropriate hardware, such as GPUs.

 ISBN 978-3-98924-119-0

A neural network can be trained on thermal imaging camera images to detect objects such as people, vehicles, or equipment. Once trained, this network will be able to automatically classify and detect objects in new frames from thermal imaging videos in real-time.

In addition to neural networks, other machine learning algorithms, such as support vector machines (SVMs), decision trees, and clustering algorithms, can also be used to process video streams in real-time. These methods may be less computationally intensive but are also effective for rapid classification or detection of objects.

Some machine learning algorithms may be less computationally intensive, allowing them to be used for real-time applications with limited resources.

Algorithms such as SVM can be efficient for tasks with a limited number of classes or low accuracy requirements.

For object classification tasks in thermal images, clustering or decision tree methods can be employed, which enable the rapid processing of data and the separation of objects from the background.

Optimizing algorithms for real-time operation is an essential component for systems that process large amounts of data, such as video streams from thermal imaging cameras. The use of specialized hardware, such as GPUs and distributed computing systems, can significantly speed up processing and reduce latency. Machine learning algorithms, particularly neural networks, are powerful tools for automating real-time object detection and classification processes. As a result, these methods allow for high accuracy and speed in processing thermal imaging data.

The development of an ontological model of thermal portraits of targets is a crucial component of modern technologies employed in military and reconnaissance systems, enabling the identification and tracking of objects using thermal cameras. Thermal portraits of targets provide accurate data on the temperature distribution of objects, allowing for the determination of their type, size, location, and other characteristics, which can be helpful for informed operational decisions.

In this context, the ontological model serves as a knowledge structure that formalizes various concepts, properties, and relationships within the domain of thermal portraits. It enables the organization of information about targets, their thermographic characteristics, and behavior, while maintaining the ability to adapt and integrate with other systems.

The ontology developed to describe thermal portraits of objects contains a number of characteristic features that ensure a structured and unambiguous representation of knowledge. It provides a formal description of the types of objects that can be captured by thermal imaging devices: technical systems and equipment, living organisms, natural objects, or special targets, each of which has its own characteristic thermal properties. For each class, thermal characteristics are determined — typical temperature ranges, maximum and minimum values, thermal radiation intensity, and other indicators that reflect the state and activity of the object and can be used for its further recognition.

An important component of ontology is the consideration of temporal dynamics. Thermal parameters are described as quantities that can change over time depending on the movement of the object, its acceleration, changes in the environment, or technical condition. This allows you to model thermal trajectories, analyze heating or cooling patterns, and predict future thermal states based on the data obtained. Spatial characteristics are also integrated into ontology: thermal indicators are linked to coordinates in two- or three-dimensional space, which makes it possible to take into account the context of the terrain, the movement of objects, and their interaction with the geographical environment.

A separate aspect is the description of external factors that affect the thermal properties of objects. Meteorological conditions such as air temperature, humidity level, or wind strength, as well as illumination and time of day, can significantly alter the thermal image. Ontology also takes into account obstacles or masking materials that can distort the thermal signal. All these elements create an integrated model that reflects the complex structure of thermal portraits and allows the construction of intelligent systems for automated processing, analysis, classification, and prediction of object behavior in thermal imaging systems.

This ontological model allows the creation of systems for automatic analysis and

 ISBN 978-3-98924-119-0

detection of threats in real time, and can also be integrated with other intelligent systems for more accurate decision-making on the battlefield or in other applications.

The main goal of this approach is to ensure high accuracy and efficiency of target detection, identification, and classification processes using thermal images. This opens up new opportunities for automated monitoring and control systems, reducing the load on the operator and increasing the speed of response in critical situations.

The task of developing a method for detecting and recognizing military equipment using video surveillance with thermal imaging channels is important for improving the safety and effectiveness of monitoring in low visibility conditions or at night. The development of a method for detecting and recognizing military equipment based on thermal imaging video is a relevant and promising area of research, as the use of such technologies can significantly improve the efficiency of monitoring and identifying objects in difficult conditions.

The main stages of this method can be as follows:

1. Data collection and preparation - a critically important stage, as the quality and diversity of video material determines the effectiveness of the future method. You can create your own database by shooting video from different angles and in different weather conditions. It is also worth using public data available through scientific papers or open sources. Data annotation must be performed to properly prepare machine learning models. This can be done manually or semi-automatically using annotation tools such as LabelImg or VGG Image Annotator. Additionally, it is important to preprocess the data, including downsampling, normalization, and increasing the amount of data using augmentation techniques such as rotation, scaling, or brightness adjustment.
2. Feature extraction - focuses on effectively extracting important features from videos that help identify and classify UAVs. Manual feature construction can include statistical features, textural features (e.g., LBP, HOG), and formal features such as aspect ratio or object compactness. Automatic feature learning using deep neural networks allows for more accurate and high-level features. For example, convolutional neural networks (CNN) are well suited for effective

feature extraction from images, while recurrent neural networks (RNN) are well suited for capturing temporal dependencies in videos.

3. Object recognition - Machine learning or deep learning algorithms, such as neural networks for classifying objects based on thermal images. Use a pre-trained model to classify objects to distinguish UAVs from other possible heat sources. Traditional methods such as k-nearest neighbors (KNN), linear discriminant analysis (LDA), or support vector machines (SVM) can be used at the classification stage. However, deep learning will be more effective for achieving greater accuracy and adaptation to complex conditions. Using CNNs with different architectures, such as VGG, ResNet, or EfficientNet, will allow for accurate object classification. Additionally, recurrent neural networks (RNNs) with LSTM or GRU cells can be utilized for video processing and analyzing temporal dependencies.
4. Analysis of results and decision-making - Evaluate the results using post-processing methods, such as spatial filters and tracking algorithms, to verify the correctness of object detection—outputting results for further processing or use in control and management systems.
5. Testing and validation - Testing the system in real-world conditions to verify its effectiveness. Adjusting system parameters based on the data obtained to improve detection and recognition accuracy.

These steps and approaches will help create a reliable and effective system for detecting and recognizing OVT based on thermal imaging video, which will be helpful for numerous applications in the field of security and military technology.

The development of algorithms for preprocessing video streams obtained from thermal imaging cameras is key to improving the effectiveness of automated surveillance systems. Specifically, key steps include data collection and preparation, noise reduction, and contrast and brightness adjustment. These stages significantly improve image quality, which in turn helps to ensure accurate object detection.

The use of various filtering methods, such as median and Gaussian filters, as well as low-frequency filtering, allows for the elimination of sporadic noise and high-frequency artifacts. Contrast correction methods, such as histogram equalization and adaptive histogram equalization, enhance the visibility of objects in images, facilitating more accurate detection of temperature anomalies, which is crucial for further analysis.

Integrating data from various sensors, including thermal imagers, optical cameras, and radars, enhances the system's accuracy and reliability in conditions of limited visibility. Interpolation and super-resolution techniques enhance image resolution, enabling more detailed and clearer thermal maps. This significantly improves the ability of surveillance systems to detect objects and anomalies in real-time.

Summary and conclusions

We analyzed the main innovative approaches to classifying thermal portraits of targets, which include the use of modern machine learning and deep learning methods, particularly neural networks such as CNN, that enable the detection of complex thermal signatures and patterns invisible to traditional methods. A crucial aspect is the application of filtering and deconvolution techniques to minimize noise and improve image quality, thereby preserving essential information about the temperature characteristics of objects. Integrating data from various sensors, including thermal cameras, radars, and satellites, enhances classification accuracy. Utilizing ontologies to structure knowledge improves interpretation and interaction with systems. In addition, multi-task learning and adaptive learning methods enable systems to continuously improve based on new data continuously, thereby significantly enhancing real-time performance.

An algorithm has been developed for preprocessing video streams obtained from thermal imaging cameras, a crucial step in processing thermographic data. Thermal imaging cameras enable the capture of images that reflect the temperature fields of objects under various conditions, which is essential to applications in military equipment, security systems, search and rescue operations, and more.

The development of methods for processing video streams from thermal imaging cameras is a crucial step in enhancing the effectiveness of automated surveillance systems, particularly in low visibility and challenging weather conditions. The use of various filtering algorithms, such as median and Gaussian filters, allows the elimination of noise and artifacts that may occur during shooting, thereby improving image quality. Contrast and brightness correction, particularly through histogram equalization and adaptive histogram equalization, enhances image clarity and facilitates more accurate detection of objects and anomalies. Integrating data from various sensors, such as thermal imagers, optical cameras, and radars, provides a more precise picture, compensating for the limitations of individual sensors and enhancing the accuracy of real-time object detection. Interpolation and super-resolution

 ISBN 978-3-98924-119-0

techniques enhance image resolution, providing additional detail for further analysis and decision-making.

 ISBN 978-3-98924-119-0

References

1. Zoriana Rybchak, Olha Kulyna, Leslav Kobyliukh. An intelligent system for speech analysis and control using customized criteria // CEUR Workshop Proceedings. – 2024. – Vol. 3723 : Proceedings of the Modern Data Science Technologies Workshop, MoDaST-2024, Lviv, Ukraine, May 31 - June 1, 2024. – P. 412–426.

2. Zoriana Rybchak, Mykhailo Kopylets. Comparative analysis of DQN and PPO algorithms in UAV obstacle avoidance 2D simulation // CEUR Workshop Proceedings. – 2024. – Vol. 3688 : Proceedings of the 8th International conference on computational linguistics and intelligent systems. Vol. III: Intelligent systems workshop. Lviv, Ukraine, Apr. 12-13, 2024. – P. 391–403.

3. Zoriana Rybchak, Olha Kulyna. Cross-cultural examination of the WAR concept: a statistical comparative analysis in English, Ukrainian, and Dutch // CEUR Workshop Proceedings. – 2024. – Vol. 3722 : Proceedings of the 8th International conference on computational linguistics and intelligent systems. Vol. 4: Computational linguistics workshop (CLW-CoLInS 2024), Lviv, Ukraine, April 12-13, 2024. – P. 557–583.

4. Zoriana Rybchak, Olha Kulyna. Synonymous variation of 'War' in the British national corpus using sketch engine: a linguistic analysis // CEUR Workshop Proceedings. – 2024. – Vol. 3723 : Proceedings of the Modern data science technologies workshop MoDaST-2024, Lviv, Ukraine, May 31 - June 1, 2024. – P. 311–327.

5. Basystiuk O., Rybchak Z., Betsa* D. Classification of military equipment based on computer vision methods // Вісник Черкаського державного технологічного університету. Технічні науки. – 2024. – том 29, вип. 3. – С. 21–30

6. Basystiuk O., Rybchak Z. Recommendation systems in e-commerce applications // Вісник Національного університету “Львівська політехніка”. Серія: Інформаційні системи та мережі. – 2024. – Вип. 15. – С. 252–259

7. Олег Басистюк, Зоряна Рибчак, Олег Буравенко*. Локалізація техніки та фортифікаційних споруд на зображенні засобами штучного інтелекту // Вісник

Хмельницького національного університету. Серія: Технічні науки. – 2024. – № 4 (339). – С. 127–129.

8. Mykich K., Zavushchak I., Savka A. Predictive maintenance for automotive vehicle engines in military logistics // CEUR Workshop Proceedings. – 2024. – Vol. 3711 : Proceedings of the Modern machine learning technologies workshop MoMLeT 2024, Lviv, Ukraine, May 31 - June 1, 2024. – P. 333–344.

9. Peleshchak I., Koshtura D., Luchkevych M., Tymchuk*** V. Classification of dynamic objects using a multilayer perceptron // CEUR Workshop Proceedings. – 2024. – Vol. 3664 : Proceedings of the 8th International conference on computational linguistics and intelligent systems. Vol. I: Machine learning workshop, Lviv, Ukraine, April 12-13, 2024. – P. 192–203.

10. Lytvyn V., Peleshchak R., Vysotska V., Peleshchak I., Chyrun L., Nazarkevych M., Vladov*** S., Lozynska O., Voloshyn S., Nagachevska O. Smart system for passive detection and classification of mines using feed-forward deep neural networks // CEUR Workshop Proceedings. – 2024. – Vol. 3861 : Proceedings of the computational intelligence application workshop (CIAW 2024), Lviv, Ukraine, October 10-12, 2024. – P. 77–111.

11. Bihun R., Karpov I. Modeling and optimization of task scheduling in multi-team IT projects // Вісник Національного університету "Львівська політехніка". Серія: Інформаційні системи та мережі. – 2024. – Вип. 16. – С. 104–111.

12. Карпов І. А., Бігун Р. Р., Буров Є. В. Використання ентропії для адаптивного коригування ваг критеріїв у багатокритеріальному виборі // Грааль науки. – 2024. – № 40 : Proceedings of the III Correspondence international scientific and practical conference "Scientific vector of various sphere' development: reality and future trends", Vinnytsia, Vienna, June 7th, 2024. – P. 347–351.

13. Карпов І. А., Бігун Р. Р. Аналіз проблем використання онтологій в системах підтримки прийняття рішень // Theoretical and practical aspects of modern scientific research : collection of scientific papers «ΛΌГОΣ» with proceedings of the IV International scientific and practical conference (Seoul, Republic of Korea, June 21, 2024). – 2024. – С. 164–165. 0,09 ум.д.ар. (За іншою тематикою) (Карпов І. А.,

Бігун Р. Р.).

14. Синько А. І. Інформаційно-технологічні інструменти формування організаційно-розпорядчих документів архіву ЗВО // Вісник Національного університету “Львівська політехніка”. Серія: Інформаційні системи та мережі. – 2024. – Вип. 15. – С. 406–418. 0,58 ум.д.ар. (За іншою тематикою) (Синько А. І.).

15. Benmhahe, B., & Chentoufi, J.A. (2021). Automated pavement distress detection, classification and measurement: A review. International Journal of Advanced Computer Science and Applications, 12(8), 708-718.

16. Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y.M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv, 2004, article number 10934. DOI: 10.48550/arXiv.2004.10934.

17. Chen, X., Liu, C., Chen, L., Zhu, X., Zhang, Y., & Wang, C. (2024). A pavement crack detection and evaluation framework for a UAV inspection system based on deep learning. Applied Sciences, 14(3), article number 1157. DOI: 10.3390/app14031157.

18. Csurka, G., Volpi, R., & Chidlovskii, B. (2022). Semantic image segmentation: Two decades of research. Foundations and Trends® in Computer Graphics and Vision, 14(1-2), 1-162. DOI: 10.1561/0600000095.

19. Figshare. (2024). Vehicles and fortification image recognition dataset. DOI: 10.6084/m9.figshare.25726470.

20. Gavrilescu, R., Zet, C., Foșalău, C., Skoczylas, M., & Cotovanu, D. (2018). Faster R-CNN: An approach to real-time object detection. In International conference and exposition on electrical and power engineering (EPE) (pp. 165-168). Iasi: IEEE. DOI: 10.1109/ICEPE.2018.8559776.

21. Hussain, M. (2023). YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. Machines, 11(7), article number 677. DOI: 10.3390/machines11070677.

22. Ippalapally, R., Mudumba, S.H., Adkay, M., & Vardhan, H.R.N. (2020). Object detection using thermal imaging. In 2020 IEEE 17th India Council international conference (INDICON) (pp. 1-6). New Delhi: IEEE India Council. DOI:

10.1109/INDICON49873.2020.9342179.

23. Iqbal, N., Mumtaz, R., Shafi, U., & Zaidi, S.M.H. (2021). Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. PeerJ Computer Science, 7, article number e536. DOI: 10.7717/peerj-cs.536.

24. Ji, G.-P., Fan, D.-P., Chou, Y.-C., Dai, D., Liniger, A., & Gool, L.C. (2023). Deep gradient learning for efficient camouflaged object detection. Machine Intelligence Research, 20, 92-108. DOI: 10.1007/s11633-022-1365-9.

25. Kishore, A.S.K., & Balamurugan, G. (2024). Faster region based convolution neural network with context iterative refinement for object detection. Measurement: Sensors, 31, article number 101025. DOI: 10.1016/j.measen.2024.101025.

26. Liang, J. (2020). Image classification based on RESNET. Journal of Physics: Conference Series, 1634, article number 012110. DOI: 10.1088/1742-6596/1634/1/012110.

27. Трофимчук О.М., Адаменко О.М., Триснюк В.М. Геоінформаційні технології захисту довкілля природно-заповідного фонду /Інститут елекомунікацій і глобального інформаційного простору НАН України; Івано-Франківськийнац. тех. ун-т нафти і газу. - Івано-Франківськ :Супрун В.П., 2021. – 343 с.// ISBN 978-617-7468-53-9 10.3997/2214-4609.201902083

28. Станкевич С. А., Масленко О. В., Андронов В. В. Адаптація нейромережевих технологій до ідентифікації малорозмірних об’єктів на супутникових зображеннях недостатньої розрізненності в базі графічних еталонів. Український журнал дистанційного зондування Землі. Київ, 2020. № 27. С. 13–17. DOI: https://doi.org/10.36023/ujrs.2020.27.175.

29. Романчук М. П. Обґрунтування типу фреймворків глибокого навчання для оброблення даних дистанційного зондування Землі. Проблеми створення, випробування, застосування та експлуатації складних інформаційних систем. Житомир, 2019. № 16. С. 70–79. DOI: https://doi.org/10.46972/2076-1546.2019.16.07.

30. Klymovych O., Hrabchak V., Lavrut O., Lavrut T., Lytvyn V., Vysotska V.

The diagnostics methods for modern communication tools in the Armed Forces of Ukraine based on neural network Approach. CEUR Workshop Proceedings. – 2020. – Vol. 2631 – P. 198–208.

31. Kuznetsova A., Rom H., Alldrin N. et al. The Open Images Dataset V4. Int J Comput Vis 128, 1956–1981 (2020). DOI: https://doi.org/10.1007/s11263-020-01316-z.

32. Wang Y. Erratum to: Towards the abstract system theory of system science for cognitive and intelligent systems. Complex Intell. Syst. 1, № 23 (2015). DOI: https://doi.org/10.1007/s40747-016-0007-7.

33. Lee J., Natsev A., Reade W. et al. The 2nd YouTube-8M Large-Scale Video Understanding Challenge. In: Leal-Taixé, L., Roth, S. (eds) Computer Vision – ECCV 2018 Workshops. ECCV 2018. Lecture Notes in Computer Science, № 11132. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-11018-5_18.

34. Benito-Gorron D., Lozano-Diez A., Toledano D. et al. Exploring convolutional, recurrent, and hybrid deep neural networks for speech and music detection in a large audio dataset. J AUDIO SPEECH MUSIC PROC. 2019, № 9. DOI: https://doi.org/10.1186/s13636-019-0152-1.

35. Godambe T., Rallabandi S., Gangashetty S.V. et al. Developing a unit selection voice given audio without corresponding text. J AUDIO SPEECH MUSIC PROC. 2016, № 6. DOI: https://doi.org/10.1186/s13636-016-0084-y.

36. Mei J., Zheng Y., Cheng M. D2ANet: Difference-aware attention network for multi-level change detection from satellite imagery. Comp. Visual Media. 2023. DOI: https://doi.org/10.1007/s41095-022-0325-1.

37. Rajpurkar P., Zhang J., Lopyrev K., Liang P. 2016. SQuAD: 100,000+ Questions for Machine Comprehension of Text. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, p. 2383–2392, Austin, Texas. DOI: 10.18653/v1/D16-1264.

38. Ramírez J., Baez M., Casati F. et al. Crowdsourced dataset to study the generation and impact of text highlighting in classification tasks. BMC Res Notes № 12, 2019. DOI: https://doi.org/10.1186/s13104-019-4858-z.

39. V. Trysnyuk, T. Trysnyuk, V. Okhariev, V. Shumeiko, A. Nikitin. Cartographic Models of Dniester River Basin Probable Flooding Centrul Universitar Nord Din Bala Mare - UTPRESS ISSN 1582-0548, №1,2018 C.61-67.

40. Mohammadi, V., Sodjinou, S.G., Katakpe, K.K., Rosse, M., & Gouton, P. (2022). Development of a multi-spectral camera for computer vision applications. In 30. jubilee international conference in central Europe on computer graphics, visualization and computer vision (pp. 24-27). Plzen: Vaclav Skala University of West Bohemia. DOI: 10.24132/CSRN.3201.4.

41. Panwar, K., Singh, A., Kukreja, S., Singh, K.K., Shakhovska, N., & Boichuk, A. (2023). Encipher GAN: An end-to-end color image encryption system using a deep generative model. Systems, 11(1), article number 36. DOI: 10.3390/systems11010036.

42. Park, S., & Lee, H. (2021). Deep learning approach to optical camera communication receiver design. In 2021 IEEE region 10 symposium (TENSYMP) (pp. 1-5). Jeju: IEEE. DOI: 10.1109/TENSYMP52854.2021.9550896.

43. Sohan, M., Sai Ram, T., & Rami Reddy, C.V. (2024). A review on YOLOv8 and its advancements. In I.J. Jacob, S. Piramuthu & P. Falkowski-Gilski (Eds.), Data intelligence and cognitive informatics (pp. 529-545). Singapore: Springer. DOI: 10.1007/978-981-99-7962-2_39.

44. Tan, M., Pang, R., & Le, Q.V. (2020). EfficientDet: Scalable and efficient object detection. In IEEE/CVF conference on computer vision and pattern recognition (CVPR) (pp. 10778-10787). Seattle: IEEE. DOI: 10.1109/CVPR42600.2020.01079.

45. Vijayakumar, A., & Vairavasundaram, S. (2024). YOLO-based object detection models: A review and its applications. Multimedia Tools and Applications, 83, 83535-83574. DOI: 10.1007/s11042-024-18872-y.

46. Yang, S., Jiang, L., Liu, Z., & Loy, C.C. (2023). GP-UNIT: Generative prior for versatile unsupervised image-to-image translation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(10), 11869-11883. DOI: 10.1109/TPAMI.2023.3284003.

47. Zhang, D., Wang, C., & Fu, Q. (2024). A new benchmark for camouflaged object detection: RGB-D camouflaged object detection dataset. Open Physics, 22(1),

article number 20240060. DOI: 10.1515/phys-2024-0060.

48. Zhang, Z., Xie, X., Guo, Q., & Xu, J. (2024). Improved YOLOv7-Tiny for object detection based on UAV aerial images. Electronics, 13(15), article number 2969. DOI: 10.3390/electronics13152969.

49. Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). Object detection with deep learning: A review. IEEE Transactions on Neural Networks and Learning Systems, 30(11), 3212-3232. DOI: 10.1109/TNNLS.2018.2876865.

50. Zhou, W., Gao, S., Zhang, L., & Lou, X. (2020). Histogram of oriented gradients feature extraction from raw bayer pattern images. IEEE Transactions on Circuits and Systems II: Express Briefs, 67(5), 946-950. DOI: 10.1109/TCSII.2020.2980557.

51. Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. Proceedings of the IEEE, 111(3), 257-276. DOI: 10.1109/JPROC.2023.3238524.

52. He, K., Zhang, X., Ren, S., & Sun, J. (2020). Deep residual learning for image recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence.

53. Rybchak, Z., & Basystiuk, O. (2017). Analysis of computer vision and image analysis techniques. ECONTECHMOD: an international quarterly journal on economics of technology and modelling processes, pp. 79-84.

54. Zou, Z., Shi, Z., Guo, Y., & Ye, J. (2021). Object detection in 20 years: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence.

55. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2021). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence.

56. Alzubaidi B., Zhang L., Humaidi J., et al. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8, 53. DOI: https://doi.org/10.1186/s40537-021-00444-8.

57. Xie, X., Cheng, G., Wang, J., Yao, X., & Han, J. (2021). Oriented R-CNN for object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence.

58. Ren, S., He, K., Girshick, R., & Sun, J. (2020). Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence.

59. Dai, J., Li, Y., He, K., & Sun, J. (2016). R-FCN: Object detection via region-based fully convolutional networks. In Proceedings of the 30th International Conference on Neural Information Processing Systems, pp. 379-387.

60. R. Lakshmi Thara, B. Upadhyay, A. Sankrityayan et al. (2024). YOLO V8: An improved real-time detection of safety equipment in different lighting scenarios on construction sites. DOI: https://doi.org/10.21203/rs.3.rs-4179998/v1.

SCIENTIFIC EDITION

MONOGRAPH

Monograph published in the author's edition

The monograph is included in
International scientometric databases

500 copies
November, 2025

Published:
ScientificWorld -NetAkhatAV
Lußstr 13,
Karlsruhe, Germany

e-mail: editor@promonograph.org
https://de.promonograph.org

ISBN 978-3-989241-19-0