UDC 358.1; 358:007.35; 534.6 DOI: https://doi.org/10.33577/2312-4458.33.2025.68-89

T. Stakh, D. Khaustov, V. Bahan, Ya. Khaustov, O. Kyrychuk, Yu. Nastishin*

Hetman Petro Sahaidachnyi National Army Academy, Lviv

METHOD FOR ESTIMATING PROBABILITIES OF TARGET DETECTION AND RECOGNITION BY THEIR DIMENSIONAL SIGNATURES IN DIGITAL IMAGES FROM THE SIGHTSEEING COMPLEX OF ARMOURED VEHICLE WEAPON SYSTEMS

The rapid development of computer vision technologies and artificial intelligence (AI) has opened up new opportunities for real-time battlefield monitoring, while simultaneously creating the challenge of efficiently selecting meaningful electronic images from massive data sets. This paper addresses the problem of automated assessment of the quality of target displaying in a digital image (QTDDI) to improve the efficiency of detecting and recognizing weapons and military equipment (WME). A system of indices of target conspicuity and salience is proposed, which enables quantitative measurement of QTDDI and the introduction of success criteria for performing visual detection and recognition tasks based on a 50% threshold probability of detection and recognition. It is claimed that only such images that meet the success criteria of visual task performance should be considered suitable for further AI-based target recognition, consequently reducing the number of false recognitions.

Special attention is paid to recognition as an intermediate stage in target data acquisition. It is proposed to employ geometric distinctive features of WME samples as their recognition signatures. An analytical expression for the recognition probability as a function of signature dimensions has been theoretically derived, and a criterion for successful performance of the visual recognition task has been established. Such an approach significantly reduces the number of false object identifications in images that do not provide the QTDDI sufficient for AI processing.

The practical implementation of the method has been demonstrated with experimental data obtained from images of the T-64BV main battle tank, collected under field conditions. Images were captured with a digital camera at various distances from the target. For these images, detection probabilities were determined based on the overall target size as the object of interest, and recognition probabilities were determined based on shape anisometry and the dimensions of characteristic signatures. The results confirmed the consistency of the theoretically developed criteria with visual analysis.

The practical significance of the work consists in establishing the methodological basis for the automated selection of sufficient-quality digital images without the involvement of an operator, which allows one to significantly reduce the amount of processed data, increase the accuracy and speed of computer vision systems, and decrease the number of false recognitions. Perspective directions for further research include extending the methodology to thermal images and using digital images for the determination of the distance to a target based on the dimensions of its signatures in electronic images.

Keywords: *target data acquisition, digital image, image processing, image fusion, artificial intelligence, visibility, conspicuity by a signature, target saliency, detection, recognition, identification, reconnaissance, unmanned aerial vehicle.*

1. Introduction and Problem Statement

With the rapid development of computer technologies and their penetration into the sphere of military applications, previously inaccessible capabilities have emerged for real-time battlefield monitoring and detection of military equipment and weapon (MEW) systems, including our armored weaponry (AW) by the enemy, as well as for detection of enemy anti-tank weapons (ATW) by our own AW using computer vision and artificial intelligence (AI) technologies.

Article history: Income 01 October 2025; Revised 02 October 2025; Accepted 03 October 2025; Primt 05 December 2025

Стах Т.М. *ORCID ID: 0000-0001-6421-9755*, Хаустов Д.Є. *ORCID ID: 0000-0001-5542-2831*,
Баган В.Р. *ORCID ID: 0000-0002-9265-8445*, Хаустов Я.Є. *ORCID ID: 0000-0003-4553-0702*
Киричук О.А. *ORCID ID: 0000-0002-1743-4978*, Настишин Ю.А. *ORCID ID: 0000-0001-7521-3906*

* Corresponding author *nastyshyn_yuriy@yahoo.com*

These modern computer technologies imply the registration of the target-background situation in the form of a digital image, at least in two spectral bands: visible and thermal (infrared).

The enormously bulky flow of TFE images obtained from multiple sources requires a tremendous amount of human labor to review, reject low-quality frames, and select meaningful video materials. Currently, this selection process – identifying frames suitable for target data acquisition (detection, recognition, and identification) – is still being performed manually by the operator.

The presence of low-quality video content can be caused by various factors such as adverse weather conditions, excessively low or blinding video/thermal signal levels, or FPV unmanned aerial vehicles leaving the target's field of view. Without prior rejection of poor-quality video materials, the AI's task of target data extraction becomes analogous to a prospector sifting vast amounts of sand in search of a few grains of gold.

In addition to substantial time consumption and computing resource load, the absence of preliminary selection of video/thermal imagery inevitably leads to a significant increase in the number of false target recognitions by AI – for instance, misidentifying buildings, buses, or trucks as armored vehicles (AV). This selection process of choosing the most significant and high-quality images should therefore be delegated to the computer.

To accomplish this, it is necessary to have methods for assessing the quality of target displaying in a digital image (QTDDI). Available image quality assessment (IQA) methods for civilian use are unsuitable for this purpose. For evaluating the QTDDI, specialized target-oriented express methods are in great demand [1–3].

The system of quantitative indices of the QTDDI developed at the Hetman Petro Sahaidachnyi National Army Academy (NASV) [4, 5] – in the form of target conspicuity indices by individual features and saliency indices by a set of distinctive features – enables quantitative evaluation of target detection performance based on saliency indices computed from image brightness tables. According to this system, an image with a higher target saliency index is considered of better quality for detection purposes.

The computation time required to determine the target conspicuity and saliency indices on a single digital image is less than 1 *ms* and can potentially be reduced to 1 μs [6], enabling evaluation of up to 10^3 – 10^6 video frames per second. The developed QTDDI index system allows for quantitative estimation of the probability of target detection (varying within the range [0,1]) in a digital image.

However, not every non-zero detection probability implies successful detection during reconnaissance. Developing quantitative criteria of detection success based on the target's distinctive features – such as its luminance-contrast signature and apparent size in the digital image – constitutes one of the key objectives of this work. As such success criteria, we propose the values of target conspicuity indices corresponding to a 50% probability of detection by a given distinctive feature.

The next stage in the target data acquisition process after detection is recognition. In this study, we introduce distinctive geometric features of the object's design – the so-called recognition signatures – which allow differentiation of a specific MEW sample from other objects. Analytical expressions for the probability of target recognition as a function of signature dimensions have been derived. Based on this analytical form, a recognition success criterion is proposed – the threshold signature size corresponding to a 50% probability of recognition of a given MEW (including AW) sample.

One of the main causes of false target recognition on electronic images by AI systems is the neglect of the fact that not every object in the image is recognizable. According to our concept, a necessary condition for successful recognition is that the probability of recognition by a given signature must exceed 50%. This threshold applies to both the size and the contrast of the signature in the image.

Before applying AI tools for object recognition on an image, it is necessary to verify that its saliency and recognition indices by the given signatures meet the criteria of successful solution of visual detection and recognition tasks. Images that fail to meet these criteria should be considered unsuitable for recognition. This will enable automatic rejection of low-quality imagery and significantly reduce the number of AI misrecognitions.

To demonstrate the practical application of the proposed approach, we conducted an experimental collection of images of the main battle tank (MBT) T-64BV at various observation distances. Using these images, we measured the probabilities of tank detection based on its overall dimensions and the probabilities of recognition based on the anisometry of its shape and the dimensions of different recognition signatures.

Aim of the Study

This study aims to derive theoretically an analytical expression for the probability of recognizing an AV in a digital image based on the characteristic geometric features (signatures) of its shape, to develop a method for estimating the probabilities of target detection and recognition by its dimensional distinctive features, and, on this basis, to establish criteria for successful accomplishing visual detection and recognition tasks in digital images as well as to verify the obtained theoretical results experimentally.

2. Method for assessing the quality of target representation based on its dimensional features of detection and recognition in digital images

2.1. Target data acquisition by digital means of target-field environment imaging.

2.1.1. Indices of target conspicuity by specific features, saliency, and concealment of the military equipment sample in a digital image. In military information sources, several approaches were reported for the quantitative determination of saliency and concealment indices of a military equipment sample. What all these approaches have in common is that the saliency/concealment of an object (a military equipment sample considered as a target) reflects the success/ failure of the target data acquisition procedure.

Accordingly, the quantitative level of target saliency or concealment is defined as the probability of success or failure of the statistical random event of collecting information about the military equipment sample as a target.

The probability of the failure event in target data acquisition is defined through the probability of its antipodal (complementary) event Q – the successful event of target data acquisition. The statistical event Q of target data acquisition is the result of three independent internal component events – detection (D), recognition (R), and identification (I) of the target, which correspond to the three stages of the target data acquisition process.

Each of these events – the overall event Q and the internal component events D, R, and I – is associated with a probability (index) of its occurrence, denoted respectively as P_Q and P_D, P_R, P_I.

The internal component events D, R, and I are independent and favorable to the overall event Q. Moreover, these events are such that the failure of any one of the three component events results in the failure of the overall event Q of target data acquisition, which is mathematically expressed as

$$P_{\mathbf{Q}}\Big|_{P_{\mathbf{D}}=0} = P_{\mathbf{Q}}\Big|_{P_{\mathbf{R}}=0} = P_{\mathbf{Q}}\Big|_{P_{\mathbf{I}}=0} = P_{\mathbf{Q}}\Big|_{P_{\mathbf{D}}=0,\, P_{\mathbf{R}}=0,\, P_{\mathbf{I}}=0} = 0. \quad (1)$$

Indeed, if the target cannot be detected ($P_D = 0$), then it cannot be recognized ($P_R = 0$) nor, much less, identified ($P_{\mathrm{I}} = 0$), which implies the failure of the overall target data acquisition event ($P_Q = 0$). If the target is detected ($P_D = 1$) but cannot be recognized ($P_R = 0$), then it also cannot be identified ($P_{\mathrm{I}} = 0$), which again implies failure of the overall target data acquisition event ($P_Q = 0$). If the target is detected ($P_D = 1$) and recognized ($P_R = 1$) – for example, as a tank – but cannot be identified as a hostile target ($P_{\mathrm{I}} = 0$), then the entire target data acquisition process is nevertheless unsuccessful ($P_Q = 0$), because these data cannot be used to make a strike decision: the inability to identify the target as hostile creates a risk of fratricide (friendly fire).

Expression (1) is a sufficient condition for classifying the events D, R, and I as being multiplicative with respect to the overall event Q of target data acquisition.

In this case, according to the methodology presented in [7], the probability P_Q of the target data acquisition event Q is equal to the product of the probabilities P_{D}, P_R, and P_I, which correspond respectively to the events of detection, recognition, and identification

$$P_{\mathbf{Q}} = P_{\mathbf{D}} \cdot P_{\mathbf{R}} \cdot P_{\mathbf{I}}. \quad (2)$$

The probability P_D of target detection depends on the method of target observation, which in turn is determined by the type of image formed as a result of observation. If the observation is carried out by the operator's eye (including with the use of optical observation devices such as binoculars or optical sights) or by an analog video camera with image output to a cathode-ray-tube (CRT) monitor, then the observation deals with an analog image. The procedure for calculating the probability of object detection on an analog image has been described in numerous monographs (see, for example, [8–10]). One of the approaches to determining the probabilities of accomplishing visual target acquisition tasks (detection, recognition, and identification of a hostile target) in an analog image is based on the methodology proposed by John Johnson [11, 12] or its more modern version – the so-called TTP model (*Targeting Task Performance*) [13]. An analysis of these models was presented in [12, 14, 15].

The essence of Johnson's criterion lies in the assumption that *visual target acquisition tasks* – namely *detection*, *recognition*, and *identification* – are considered successfully accomplished if the probability of solving the corresponding visual task exceeds 50% (0.5 on the probability scale [0;1]). The same criterion is followed within the framework of the TTP model (Targeting Task Performance). To apply this criterion, a target saliency criterion on the image is introduced. In Johnson's approach, the saliency of the target is represented by the apparent size of the target on the image, expressed in terms of the number N of line-pair divisions of a reticle that cover the target image. The number of line pairs corresponding to a 50% probability of successfully solving a given visual task is denoted as N_{50}^{task}. In the TTP model, the corresponding integral saliency indices, denoted V and V_{50}^{task}, are introduced. The target saliency index V is defined by a specific expression representing the integral of the functions of target contrast and size. The probability of successfully accomplishing the corresponding visual target acquisition task is then determined by the following expression [16, 17]

$$P_{task}(x) = \frac{\left(x/x_{50}^{task}\right)^{y}}{1+\left(x/x_{50}^{task}\right)^{y}}, \quad (3)$$

where the variables x, x_{50}^{task} take the corresponding values $x = N$, $x_{50}^{task} = N_{50}^{task}$ and $y = 2.7 + 0.7\left(N/N_{50}^{task}\right)$ in the Johnson model, and $x = v$, $x_{50}^{task} = v_{50}^{task}$ $y = 1.51 + 0.24\left(v/v_{50}^{task}\right)$ – in the TTP model. The quantities N_{50}^{task} and v_{50}^{task} are considered to be tabulated values. It was shown experimentally in [16] and proved theoretically in [17] that the variables in the Johnson and TTP models are related by the following relationship

$$v = \chi N - v^{(0)}. \quad (4)$$

When performing calculations for a specific visual task, the superscript of the *task* in expression (3) should be replaced by the index corresponding to the respective visual task – *detection* (*D*), *recognition* (*R*), or *identification* (*I*). According to the literature data [17] the resolution thresholds for *detection* N_{50}^{D}, *recognition* N_{50}^{R}, and *identification* N_{50}^{I} are related as $N_{50}^{D} : N_{50}^{R} : N_{50}^{I} = 1:4:8$. It is important to note that this ratio $N_{50}^{D} : N_{50}^{R} : N_{50}^{I} = 1:4:8$ numerically coincides with the ratio $\varepsilon_D : \varepsilon_R : \varepsilon_I = 1:4:8$ of the threshold contrast values for the human visual system [16]. This indicates that, although *de jure* in Johnson's methodology the parameter N_{50}^{task} is ascribed the physical meaning of the target size in the image, *de facto* it is also indirectly related to the target's contrast in the image. The reason is that the quantities N_{50}^{task} are determined experimentally, based on expert visual observations, where observers subconsciously react not only to the target's apparent size but also to its contrast. Therefore, it is not surprising that the same ratios $N_{50}^{D} : N_{50}^{R} : N_{50}^{I} = 1:4:8$ were found to hold both for the threshold contrast values $\varepsilon_D : \varepsilon_R : \varepsilon_I = 1:4:8$ and for the resolution parameters $v_{50}^{D} : v_{50}^{R} : v_{50}^{I} = 1:4:8$.

Indeed, the experimental data obtained in [16] indicate that the value of v_{50}^{task} can be considered related to N_{50}^{task} by the approximate relation $v_{50}^{task} \approx 2.7 N_{50}^{task}$ for all three visual tasks – detection, recognition, and identification. From this it follows that, for the resolution thresholds of *detection* v_{50}^{D}, *recognition* v_{50}^{R}, and *identification* v_{50}^{I} within the framework of the TTP model, the same ratio also holds $v_{50}^{D} : v_{50}^{R} : v_{50}^{I} = 1:4:8$.

The presence of the common variable x in expressions (3) for all visual tasks (detection, recognition, and identification), together with the fulfillment of relations

$$x_{50}^{D} : x_{50}^{R} : x_{50}^{I} = N_{50}^{D} : N_{50}^{R} : N_{50}^{I} = v_{50}^{D} : v_{50}^{R} : v_{50}^{I} = 1:4:8 \quad (5)$$

may suggest that the probabilities of detection $P_{\mathbf{D}}(x)$, recognition $P_{\mathbf{R}}(x)$, and identification $P_{\mathbf{I}}(x)$ are not *independent functions*. This, in turn, would mean that the statistical events *D*, *R*, and *I* are not independent. Intuitively, it may seem that the higher the probability of detection, the higher one can expect the probability of recognition to be; and the higher the detection and recognition probabilities, the greater – at first glance – appears to be the probability of target identification. Such a relationship could imply that the probabilities $P_{\mathbf{D}}(x)$, $P_{\mathbf{R}}(x)$, and $P_{\mathbf{I}}(x)$ are dependent functions.

In that case, the analytical dependence of the functions $P_{\mathbf{D}}(x)$, $P_{\mathbf{R}}(x)$ and $P_{\mathbf{I}}(x)$, given by expression (3) could imply that expressions (3) and (5) form a system of equations. If this were true, then by eliminating the common variable from this system of equations, one could expect to analytically express the probability of accomplishing one visual task through the probability of accomplishing another visual task. The possibility of analytically expressing one visual task's probability in terms of another would cast doubt either on the validity of expressions (3) and (5) themselves, or on the assumption that the events of detection, recognition, and identification are statistically independent.

Verification of the assumption that the events of detection (*D*), recognition (*R*), and identification (*I*) are independent is important, since the existence of a functional relationship between different visual tasks contradicts the generally accepted premise that the statistical random events of detection, recognition, and identification are independent.

For example, if the numerical value of the target *detection* probability $P_{\mathbf{D}}$ on an electronic image was obtained as a result of measuring the target saliency index Γ for a specific image, one might expect that the probabilities of *recognition* $P_{\mathbf{R}}$ and *identification* $P_{\mathbf{I}}$ could be determined through the known value of $P_{\mathbf{D}}$, taking into account the relationship (5), according to which $x_{50}^{D} : x_{50}^{R} : x_{50}^{I} = 1:4:8$. However, such an approach is impossible for several reasons, and there is no contradiction between expressions (3), (5), and the statement that the events of detection, recognition, and identification are independent. The arguments supporting this statement are presented below.

It should be noted that the term *dependence/ independence of functions* has a broader meaning than the term linear *dependence/independence of functions*.

Two functions (x) and g(x) sharing a common variable x may be *linearly independent* but still might exhibit a *nonlinear dependence.* A rigorous mathematical proof of the independence of the functions $P_{\mathbf{D}}(x)$, $P_{\mathbf{R}}(x)$ and $P_{\mathbf{I}}(x)$, defined by expression (3), goes beyond the scope and objectives of this paper; therefore, it will be the subject of our subsequent publication. Here, however, we briefly present several arguments that support the mathematical *independence* of the probabilities corresponding to different visual tasks.

Firstly, the presence of a common variable x in the dependencies $P_{\mathbf{D}}(x)$, $P_{\mathbf{R}}(x)$ and $P_{\mathbf{I}}(x)$ is only a necessary condition for their functional dependence, but not a sufficient one. There exist analytical functions of a common argument that remain independent. For example, the functions $f(x)=x$ and $g(x)=\sin(x)$ are independent, even though both are functions of the same variable x.

Secondly, one of the ways to test functions with a common argument for dependence is to construct the inverse dependencies of the argument with respect to the functions, that is, $x(P_{\mathbf{D}})$, $x(P_{\mathbf{R}})$, and $x(P_{\mathbf{I}})$

Since x is the common argument for all three functions, the relationship $x(P_{\mathbf{D}})=x(P_{\mathbf{R}})=x(P_{\mathbf{I}})$, should hold, from which the dependencies $P_{\mathbf{R}}(P_{\mathbf{D}})$, $P_{\mathbf{I}}(P_{\mathbf{D}})$ or $P_{\mathbf{I}}(P_{\mathbf{R}})$ could, in principle, be determined. However, according to formula (3), the dependencies $P_{\mathbf{D}}(x)$, $P_{\mathbf{R}}(x)$ and $P_{\mathbf{I}}(x)$ are transcendental sigmoid functions, and therefore, it is impossible to express the inverse functions $x(P_{\mathbf{D}})$, $x(P_{\mathbf{R}})$, and $x(P_{\mathbf{I}})$ in analytical quadratures. The impossibility of expressing the common argument x of the functions $f(x)$ and $g(x)$ in terms of the functions f and g themselves within analytical quadratures is an indication that the functions $f(x)$ and $g(x)$ are *independent*.

Thirdly, in order to determine, for example, the dependence $P_{\mathbf{R}}(P_{\mathbf{D}})$ by eliminating the common variable x, the values of the function $x(P_{\mathbf{D}})$ must be evaluated at those values of $x > x_{50}^{R}$ for which, according to the Johnson criterion, the visual task of recognition is considered successful. Taking into account relation (5), this condition can be written as $x > 4x_{50}^{D}$. Since the probabilities $P_{task}(x)$ are sigmoid functions, in the case when $x > 4x_{50}^{\mathbf{D}}$, the detection probability $P_{\mathbf{D}}\left(4x_{50}^{\mathbf{D}}\right)=4^{y}/\left(1+4^{y}\right)$ approaches saturation, so that $P_{\mathbf{D}} \to 1$. In the Johnson model, $y\left(4N_{50}^{D}\right)=2.7+0.7\left(N/N_{50}^{задачі}\right)=5.5$, and therefore, as, we obtain $P_{\mathbf{D}}\left(4x_{50}^{\mathbf{D}}\right)=0.9995$. For large values of $x > 4x_{50}^{\mathbf{D}}$, as the argument x increases, the computational error in determining the value of $x(P_{\mathbf{D}})$ from expression (3) tends to infinity. Indeed, the error in determining x from the function $P_{\mathbf{D}}(x)$ can be expressed as follows

$$\Delta x\Big|_{x>4x_{50}^{\mathbf{D}}}=\frac{\Delta P_D}{\left(\dfrac{dP_D}{dx}\right)\Big|_{x>4x_{50}^{\mathbf{D}}}}, \qquad (6)$$

where the numerator ΔP_D – the error in determining the detection probability as the target saliency in a digital image according to expression (8) – is a nonzero finite number within the range $]0;1[$, and the denominator represents the derivative of the detection probability with respect to its argument at a given value of $x > 4x_{50}^{\mathbf{D}}$. For the lower limit of the Johnson criterion $N = 4N_{50}^{\mathbf{D}}$, within the Johnson model we obtain a relative error in determining the parameter N on the order of $2\times10^{4}\%$, which increases following a power-law dependence as the value of N grows. For the lower threshold of the criterion corresponding to the successful accomplishment of the identification task $N = 8N_{50}^{\mathbf{D}}$, the relative error is on the order of $2\times10^{8}\%$. Thus, for large values of $x > 4x_{50}^{\mathbf{D}}$, as x increases, the value of the derivative in the denominator of expression (6) $\left(\dfrac{dP_D}{dx}\right)\Big|_{x>4x_{50}^{\mathbf{D}}} \to 0$, and consequently, the error in calculating the value of $x(P_{\mathbf{D}})$ from expression (3) tends to infinity. A value of $x(P_{\mathbf{D}})$ determined with infinite error, when substituted into the dependence $P_{\mathbf{R}}(x(P_{\mathbf{D}}))$, results in an infinite error in determining the recognition probability $P_{\mathbf{R}}(P_{\mathbf{D}})$ as a function of the detection probability $P_{\mathbf{D}}$. An infinite error in determining $P_{\mathbf{R}}(P_{\mathbf{D}})$ and $P_{\mathbf{I}}(P_{\mathbf{D}})$ therefore indicates that, in reality, there is no functional relationship between the probabilities $P_{\mathbf{D}}$, $P_{\mathbf{R}}$ and $P_{\mathbf{I}}$, and hence, that the events of detection (D), recognition (R), and identification (I) are independent.

Fourthly, another argument in favor of the independence of the events of detection, recognition, and identification is that, according to definition (3), the probabilities of all these events take on the same value when the corresponding argument assumes the same value

$$P_{task}\big|_{\forall x_{50}^{task}} = \frac{1}{2} \tag{7}$$

or all values of N_{50}^{task} or v_{50}^{task}, despite the fact that a functional relationship (5) exists between the quantities N_{50}^{task}, and v_{50}^{task}. Thus, under the imposed condition (7), the entire range of values of x_{50}^{task} corresponding to different visual tasks (detection, recognition, and identification) is projected onto a single point $P_{task} = 1/2$. In other words, a single value of the probability $P_{task} = 1/2$ corresponds to a nontrivial domain (set) of values of x_{50}^{task} for different visual tasks. In turn, this means that there is no functional relationship between the quantities corresponding to different visual tasks, i.e., the functions $P_{\mathbf{D}}$, $P_{\mathbf{R}}$ and $P_{\mathbf{I}}$ are independent, and therefore the statistical events *D*, *R*, and *I* are independent.

The probability P_D of detecting a target in a digital image defines the target's saliency index *Г* for that image. Note that hereafter in the text the term *target* denotes an *object of interest* in the image, as commonly used in computer vision theory – a notion broader than the term *military target*. On the battlefield, an object of interest may be a *potential* military target. An object of interest becomes a military target only after the data-collection process for that object has been successfully completed, i.e., when the object has been detected, recognized and identified as hostile and therefore warranting engagement.

The saliency index *Г* of a target in a digital image is defined as the probability of detecting the target based on the set of conspicuity features, and is computed from the individual conspicuity indices associated with each feature

$$\Gamma = V \cdot P_S \cdot P_{clutter}, \tag{8}$$

where *V* is the target visibility index, which in essence represents the conspicuity index based on luminance contrast; P_s is the conspicuity index based on the target's size in the digital image; and $P_{clutter}$ is the conspicuity index of the target against a cluttered background. The procedure for determining the conspicuity indices and the overall saliency index is described in detail in works [4–7, 18].

It is important to note that in the model of the QTDDI, the size-based conspicuity index P_s is an analogue of the probability $P_{task}\left(N/N_{50}^{task}\right)$, which is determined by the target size *N* on an analog image in the Johnson model, as given by expression (3). However, the experimental determination of the parameter N_{50}^{task} through visual observation and expert evaluation imposes an implicitly (hidden) expressed functional dependence of this value on the local contrast of the target. Therefore, in practice, $P_{task}\left(N/N_{50}^{task}\right)$ represents an implicit combination of the target visibility index *V* and the size-based conspicuity index P_s.

In the TTP model, the dependence $P_{task}\left(\dfrac{v}{v_{50}^{task}}\right)$ explicitly incorporates the parameter's dependence on both the target's local contrast and its size. Thus, the probability $P_{task}\left(\dfrac{v}{v_{50}^{task}}\right)$ in that model corresponds to a combination of the indices *V* and P_s. The advantage of our QTR-EI model is that the indices *V* and P_s are treated as independent functions, whereas in the TTP model they are functionally combined under the integral sign. This combination makes it impossible to determine separately the individual contributions of conspicuity due to contrast and conspicuity due to size.

The saliency *Г* of an armored vehicle (AFV) sample in an electronic image can be evaluated through experimental measurements performed on images obtained from the digital cameras of the tank's sighting and observation complex (SOC), as well as from cameras mounted on an unmanned aerial vehicle (UAV), using expression (8).

3. Criterion for successful accomplishment of the visual task of target detection by a conspicuity feature

For analog images, the Johnson criterion has proven to be an effective measure, according to which the visual task of target data acquisition is considered successfully accomplished if its probability exceeds 50%. Similarly, in the QTDDI model, parameters can be introduced for the conspicuity indices corresponding to different features, which represent a 50% probability of accomplishing the visual task based on a given conspicuity feature.

3.1. Criterion for successful target detection by the luminance contrast feature.

Similar to the Johnson criterion, the task of target detection by the luminance contrast feature is considered successfully accomplished if the target visibility index

$$V > V_{50} = \frac{1}{2}. \tag{9}$$

According to definitions [15, 16], the target visibility index is determined by the expression

$$V = \begin{cases} \dfrac{|K| - |K_{th}|}{|K| + |K_{th}|\left(1 - 2|K|\right)} & at\ |K| \ge |K_{th}|, \\ 0\ at\ |K| < |K_{th}| \end{cases} \tag{10}$$

where $|K|$ is the absolute value of the target's local contrast, and $|K_{th}|$ is the threshold value of local contrast that can still (or just) be distinguished by the human eye.

By substituting expression (10) into expression (9), we obtain an expression for the normalized local contrast $|K_{50}|$, which corresponds to a 50% probability of target detection by the luminance contrast feature

$$|K_{50}| = \frac{3|K_{th}|}{1+2|K_{th}|}, \quad (11)$$

and consequently, the expression for the criterion of successful target detection by the luminance contrast feature

$$|K| > |K_{50}| = \frac{3|K_{th}|}{1+2|K_{th}|}. \quad (12)$$

The value of the threshold contrast $|K_{th}|$ is determined by the so-called Fisher threshold *p*-level $|K_{th}| = 0.05$ [19, 20]. By substituting the value of $|K_{th}| = 0.05$ into expression (12), we find that the visual task of target detection by the luminance contrast feature is considered successfully accomplished if the normalized local contrast of the target satisfies the condition

$$|K| > 0.14. \quad (13)$$

3.2. Probability and criterion for successful target detection by the size feature in the image.

Detection of a target as an object of interest in a digital image is reduced to the perception of the object as a whole; therefore, its 2D size – the area S – is the key dimensional parameter in the case of target *detection*. For this reason, within the framework of the QTDDI model, the size-based conspicuity index for target *detection* is defined as a Bayesian *geometric* probability [4, 5].

$$P_s = \frac{\frac{s}{S}\left(1-\frac{s_{th}}{S}\right)-\frac{s_{th}}{S}\left(1-\frac{s}{S}\right)}{\frac{s}{S}\left(1-\frac{s_{th}}{S}\right)+\frac{s_{th}}{S}\left(1-\frac{s}{S}\right)}, \quad (14)$$

where S, s and s_{th} are the 2D sizes, the areas of the target, the target expectation zone (TEZ) [21], and the minimum threshold image element area that can still be distinguished by size, either by a human observer or a computer, respectively. When target detection is performed by a human operator, $s_{th} = l_{th}^2$ corresponds to the area of a square subtending a solid angle of one square arcminute, which represents the average angular resolution of the human eye. The minimum distance between two points that can be resolved by the human eye at a viewing distance of 50*cm* from the monitor, corresponding to the average angular resolution of one arcminute, is l_{th} = 0.145*mm*. Hence, for target detection by an operator on an image displayed on a computer monitor, the condition is $l > 0.145mm$. On the image plane, the average visual resolution of one arcminute corresponds to an area of $s_{th} = l_{th}^2 = 0.145^2(mm^2) = 0.021(mm^2)$. For a 13.3-inch Retina laptop monitor with a resolution of 2560×1600 *pixels*, the threshold length l_{th} = 0.145*mm* is approximately comparable to the size of one pixel (≈0.112*mm*). For monitors with significantly lower resolutions, for example 800×600 *pixels*, the pixel size is larger and equals 0.254*mm* for a 10-inch diagonal, 0.381*mm* for a 15-inch diagonal, and 0.432*mm* for a 17-inch diagonal. Since a digital image is displayed on a monitor, the visual resolution of the human eye cannot be better than one pixel. In addition, according to Nyquist [22, 23], the maximum spatial resolution frequency (1 line pair per millimeter), which corresponds to the minimum resolvable spatial period of the monitor (or a video/thermal camera), equals l_{th}=2*pixels* (*px*). Accordingly, the visual resolution of the eye when observing an image on a monitor should not be assumed to be better than two pixels. Therefore, in the following, we assume that l_{th} = 2*px* and $s_{th} = l_{th}^2 = 4\left(px^2\right)$.

Before proceeding, it is necessary to clarify the meaning of the terms TEZ and *geometric* probability. The TEZ refers to a strip (region) within the image where it makes sense to search for the target. The reasoning is that a significant portion of an image may correspond to areas of space where the target definitely cannot be located. For example, it is unreasonable to search for an armored vehicle in the sky, on a water surface, in tree canopies, or in dense forest areas. Therefore, TEZ is typically much narrower than the entire image frame. In most cases, the vertical dimension – the height of the TEZ *H* – is smaller than the image height, while its length may equal the image frame length *L*, or it may be smaller $L_s < L$ if there are reasons to restrict the search area on one or both sides of the image. For instance, this may occur when a road along which armored vehicles move passes through a forest, or when there is deep water on both sides of the road.

As for the term *geometric probability*, according to definition [24], it is the probability that a random point will fall within a certain region when that region is defined by a geometric figure. In the case of target search within an image, the "*falling of a random point*" is understood as *the random gaze of an observer falling* within the area of interest (the target) with an area *s*, which is located within TEZ of area $S = L_s \times H_s$ in the image. Therefore, the ratio *s/S* is nothing else but the geometric probability of detecting a target with area s located within the TEZ background of area S. The quantity s_{th} represents the area of the forbidden zone, into which the observer's gaze cannot fall on the image background. Hence, the ratio s_{th}/S corresponds to the probability of non-detection of the target due to the nonzero value of the threshold area s_{th}. With the condition $s_{th} \neq 0$ imposed, the probability of target detection in an electronic image reduces to a Bayesian probability, given by equation (14), where – up to a common coefficient – the numerator represents the conditional probability, and the denominator represents the total probability of detecting a target of area s within a TEZ of area S, under the condition $s_{th} \neq 0$.

After simplification, expression (14) can be rewritten in the form

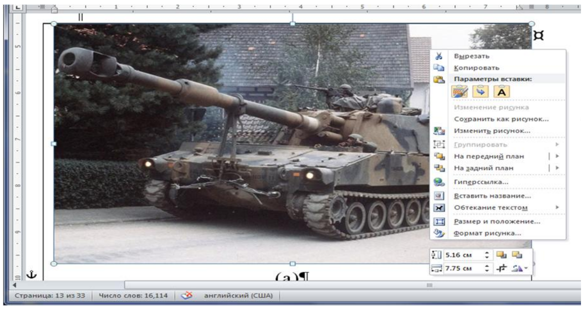
$$P_s = \begin{cases} \dfrac{s(S-s_{th})-s_{th}(S-s)}{s(S-s_{th})+s_{th}(S-s)} & at \quad s_{th} \le s \le S \\ 0 \quad at \quad 0 \le s \le s_{th} \end{cases} \tag{15}$$

or, in a more compact form, as

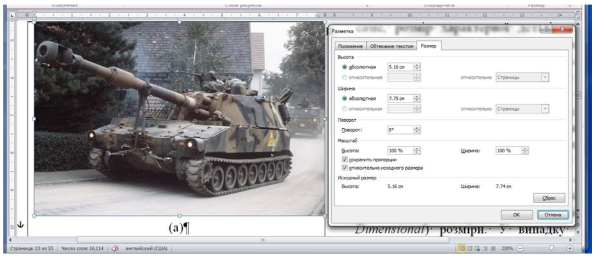
$$P_s = \begin{cases} \dfrac{1-\dfrac{s_{th}}{s}}{1+\dfrac{s_{th}}{s}\left(1-2\dfrac{s}{S}\right)} & at \quad s_{th} \le s \le S \\ 0 \quad at \quad 0 \le s \le s_{th} \end{cases}. \tag{16}$$

The representation in the form of (16) is convenient because it can be easily analyzed both in units of ordinary length (e.g., in mm^2) and in square pixels (px^2), since all variables enter the formula as dimensionless ratios s_{th}/s and s/S.

The areas of the TEZ S and the target s for a given electronic image can be determined from the digital data of the image itself. For example, if the monitor resolution is 800×600*pixels*, the image occupies the entire screen (a full-screen image), and the TEZ covers the entire image, then the horizontal dimension (width or length) of the TEZ is L=800*pixels*, and the vertical dimension (height) is H=600*pixels*. In that case, the area of the TEZ is $S=L\times H=800\times600=480{,}000\,px^2$. If the image does not occupy the entire screen but is embedded in a document – for example, in Microsoft Word (MW) – then, to determine the image size, one needs to right-click on the image and read its width (L) and height (H) from the bottom part of the drop-down menu (Fig. 1a). Alternatively, one can select "Size and Position" from the drop-down menu, and in the window that appears, click on the "Size" tab (Fig. 1b) to view the corresponding dimensions.



a



b

Fig. 1. **Reading the image size: a from the drop-down menu; b – from the "Size" window**

As a rule, in MW the image size is displayed by default in centimeters or inches. The image size in pixels can be calculated using the formula

$$Number\ of\ pixels = \frac{Image\ size\ (in\ cm)}{2{,}54} \times \mathrm{DPI}, \tag{17}$$

where it is taken into account that 1 inch=2.54*cm*, and DPI (*dots per inch*) represents the image resolution. By default, in MW the value is DPI=96. (One can verify this by right-clicking on the inserted image file in MW and selecting Properties → Details from the drop-down menu.).

It is important to emphasize that formulas (14)–(16) are intended for calculating the probability of target detection by its size feature specifically in a digital image. A distinctive advantage of digital electronic imagery is that it can be rescaled, increasing the image frame size to an area S convenient for visual observation. Moreover, using the "zoom" function, the image can be enlarged so that the target appears magnified. The zoom procedure consists of cropping the image so that the target, as the object of interest, is centered in the frame, and then stretching the cropped image so that its frame coincides with the frame of the original uncropped image (i.e., "mapping" the cropped image onto the original frame). This mapping occurs due to an increase in the linear pixel size. Equivalently, the zoom procedure can be interpreted as an increase in the linear dimensions of the pixel, which results in an enlargement of the image frame itself; therefore, the frame of the enlarged image is cropped again along the contour of the original (pre-zoom) frame.

If the target size s is significantly larger than s_{th}, the zoom procedure can be used to enlarge the target size s until it becomes comparable to the image size S; in this case, the size-based conspicuity index of the target $P_s|_{s\to S} \to 1$. In other words, if $s \approx s_{th}$, the zoom operation allows the target size in the image to be adjusted so that $s \approx S$ and therefore, according to expression (15), we obtain $P_s|_{s\gg s_{th}} \to 1$.

The zoom procedure changes the target size s in millimeters, while the size of the image frame remains unchanged. That is, due to the zoom operation, the variable in expressions (15) and (16) is the target size s, whereas S=const if the dimensions are measured in units of length (e.g., mm^2). As a result, we arrive at an evident conclusion: during the zoom procedure, an increase in s leads to an increase in the ratio s/S, and consequently, to an increase in the probability of target detection P_s.

If the analysis is carried out in units of square pixels (px^2), then when the target size increases in conventional metric units of length (e.g., mm^2), the number of px^2 corresponding to the target remains unchanged, i.e., $s[px^2]$=const, while, due to the cropping of the image to match the frame size of the original (non-zoomed) image, the total number of pixels corresponding to the entire image $S[px^2]$ decreases by the same factor by which the target size increased in square length units. As a result, we again arrive at a logical conclusion: the zoom procedure leads to an increase in the ratio s/S, and consequently, to an increase in the probability of target detection P_s.

Thus, the higher the image resolution, the greater the number of pixels that correspond both to the entire image and to the target. The resolution of an image is measured by the number of pixels along the horizontal and vertical axes. For example, if the monitor resolution increases from 800×600 to 1320×990, the number of pixels s corresponding to the target increases by a factor of 2.73. The same applies to the image resolution measured in DPI (dots per inch). Therefore, *the higher the monitor resolution* and *image* DPI, *the higher the probability of target detection.* Moreover, higher monitor and image resolution values mean a larger number of pixels per target, while the threshold resolution remains constant at $s_{th} = 4px^2$. According to expression (16), this also leads to an increase in the size-based conspicuity index of the target.

It is important to note that due to the nonzero value $s_{th} \neq 0$, zooming (magnification) cannot be performed indefinitely. Therefore, for values of S close to s_{th}, the question arises:

What should be the value of the target size relative to $s_{th} \neq 0$ for the task of target detection by the size feature on an electronic image to be considered successfully accomplished?

Similar to the criterion for successful accomplishment of the target detection task by the luminance contrast feature, we introduce the criterion for successful accomplishment of the target detection task by the size feature on in a digital image

$$P_s > (P_s)_{50} = \frac{1}{2}. \tag{18}$$

From expressions (18) and (15) (equivalently, (16)), we obtain the value of the target size in the digital image that ensures a 50% probability of target detection

$$s_{50} = \frac{3s_{th}}{1+2\dfrac{s_{th}}{S}}. \tag{19}$$

When using formula (19), it should be highlighted that for a nonzero $P_s \neq 0$, the condition $s > s_{th}$ must be satisfied, and therefore the condition $s_{th} < s < S$ must also hold. The condition $s \gg s_{th}$ is practically always satisfied, and therefore

$$S_{50}|_{S\gg s_{th}} = 3s_{th}. \tag{20}$$

Considering that $s_{th} = 4px^2$, we arrive at the conclusion that the condition for the successful accomplishment of the visual task of target detection can be written in the form

$$s > 12\,px^2. \tag{21}$$

In other words, for successful detection in the image, the target must have at least the shape of a rectangle with sides of $4\times3px^2$ or a square with a side of at least $4px$.

It is evident that increasing the resolution of both the monitor and the image enhances the successful accomplishment of the visual task of target detection by the size feature on an electronic image.

4. Probability of target recognition by its characteristic dimensional features in a digital image

4.1. Characteristic recognition features and the probability of target recognition by a specific conspicuity feature.

4.1.1. Recognition features. The success of recognizing an object in a digital image depends on the conspicuity of its recognition features. The conspicuity features used for target detection – such as target visibility (luminance contrast), image size, and conspicuity against a cluttered background – are not recognition features. In essence, the process of target recognition also reduces to a detection process, but not of the target as a whole; rather, it concerns the detection of a distinct geometric characteristic or structural element that makes it possible to distinguish, for instance, a combat armored vehicle (CAV) from other types of vehicles. For a tank, such characteristic recognition features include the turret with the gun barrel and the tracks. Therefore, target recognition in an electronic image is based on a separate and independent set of recognition features. It is precisely this property of an independent set of recognition features that determines the independence of the recognition (*R*) and detection (*D*) events. In military literature, *the recognition features of a target* are referred to as *signatures*, which is the Ukrainian transliteration of the English term signature, literally meaning "*personal signature*" – in the sense of a distinctive or *characteristic feature*.

The most important *recognition feature* (*signature*) of a target is *the characteristic shape of the vehicle's contours*. *Shape signatures* such as the elongated form of the gun barrel and the distinctive outline of the turret make it possible to recognize a combat vehicle as a tank. However, these features alone are not always sufficient to distinguish a tank from a self-propelled howitzer on a tracked chassis. For example, a self-propelled artillery system (SPH) such as the M109 howitzer (Fig. 2) in its various modifications – M109A2 (a), M109A7 (b), and M109L with anti-drone protection (c) – resembles a tank in overall shape. An auxiliary distinguishing feature that differentiates an SPH from a tank is *anisometry of form*, meaning the difference in dimensions along different spatial directions.

Firstly, the height-to-length ratio for a tank is noticeably smaller than that for a SPH. Secondly, as a rule (though not always), the barrel length of an SPH is greater than that of a tank, so that the ratio of barrel length to vehicle length is higher for an SPH compared to a tank. Tanks are designed for close combat, infantry support, and breakthrough of the enemy's front line; therefore, they rarely operate individually and usually act within a tank unit, at least in pairs. In contrast, SPHs are intended for long-range fire support, destruction of fortifications and enemy artillery positions, and providing covering fire, and thus operate at a distance from direct engagement zones, using artillery positions to support friendly forces. Hence, *the nature of combat employment* also serves as a recognition feature.

Moreover, the nature of combat operations of AFV leaves a clear imprint on their design characteristics. Because tanks are required to perform missions demanding mobility, maneuverability, and stealth, the turret of a tank is noticeably smaller and lower than that of a SPH. The superstructure above the tracks is also lower in tanks. In contrast, both the turret and superstructure of SPHs are higher and bulkier due to the larger dimensions of the artillery installation they accommodate.

Thus, at least the following recognition features of CAVs can be distinguished, listed in order of their significance:

1. Presence and characteristic shape of specific components:
   • gun barrel,
   • turret,
   • tracks.
2. Anisometry of the hull shape: ratios of
   • vehicle height to its length,
   • track height (road wheel diameter) to total vehicle height,
   • turret length to overall vehicle length.
3. Nature of combat employment.
4. Sound of the running engine.

a

B

c

Fig. 2. **Self-propelled howitzer: a – M109A2; b – M109A7; c – M109L with anti-drone protection [25]**

Only the first two recognition features listed above are geometric conspicuity features of target recognition in an electronic image.

Both of these features reduce to the size-based conspicuity index, but unlike *the conspicuity* features used for detection, in this case we are dealing not with the size of the target as a whole, but with the size of a characteristic shape or detail of the vehicle that enables its recognition – for example, the gun barrel or the turret, identified by their distinctive shapes or by characteristic differences in overall dimensions, such as the anisometry of form.

In particular, the T-64BV tank has the following dimensions: length l=9m and height h=2.16m. Hence, h/l=0.24. For the M109A2 self-propelled howitzer, l=9 m, h=3.25m, and thus h/l=0.36. Therefore, the ratio of vehicle height to its length (h/l) could serve as one of the recognition features (signatures) for distinguishing between a tank and an SPH. However, due to the mobility of the vehicle, namely its ability to rotate around its vertical axis, this signature can only be used effectively when the vehicle is observed side-on (perpendicularly). For example, in Figures 2a-c, the self-propelled howitzers are viewed partially from the side. As a result, instead of the true vehicle length l, only the projected length $l_{pr}=l\ sin\beta$ on the image plane can be measured, where β is the angle between the line of sight and the longitudinal axis of the vehicle.

Since the value of the projection may vary within the range $l_{pr} \in [0;1]$ depending on the angle β formed by the normal to the image plane and the side plane of the vehicle (the viewing angle), it is clear that the parameter h/l_{pr} cannot always be used to reliably distinguish a tank from a SPH.

The problem with using the ratio h/l for vehicle recognition lies in the fact that only the value of h is *invariant* (remains unchanged) with respect to rotations around the vertical axis, whereas on the image we measure not l itself but its projection l_{pr}, which is not *invariant* with respect to such rotations. Hence, we are led to conclude that *for successful vehicle recognition, it is necessary to use the geometric lengths of structural elements that are invariant with respect to the same axis – whether vertical or horizontal.*

Such a signature could be the ratio of the total vehicle height h to the height of the tracked running gear, i.e., the diameter of the road wheels d_k. In this case, both quantities – d_k and h are invariant with respect to rotations around the vertical axis.

From Table 1 it can be seen that for the T-72, T-90, and Challenger 2 tanks, the ratio h/d_k≈3, whereas for the self-propelled howitzers “Akatsiya” and “MSTA-S”, h/d_k≈5. Thus, as a rule, SPHs have a higher superstructure together with the turret above the track system compared to tanks, which is due to the significantly larger dimensions of the SPH’s artillery installation.

Table 1

Road wheel diameter d_k, vehicle height h, and their ratio for tanks and SPHs

Vehicle model	Height h, m	Road wheel diameter, $d_к$, m	Ratio $h/d_к$
T-64(tank)	2.17	0.55	3.8
T-80(tank)	2.20	0.67	3.2
T-72(tank)	2.23	0.75	2.9
T-90 (tank)	2.22	0.75	2.9
Challenger 2 (tank)	2.49	0.70	3.5
2C1 "Gvozdika" (SPH)	2.28	0.67	3.4
2C3 "Akatsiya" (SPH)	3.05	0.63	4.8
2C19 "MSTA-S" (SPH)	3.30	0.52	5.0
2C35 "Koalitsiya-SV" (SPH)	3.5	0.75	4.7
2C5 "Giatsint" (SPH)	2,76	0,63	4.3
2C4 "Tyulpan" (SPH)	3.25	0.63	5.1
M110 (SPH) (USA)	3.28	0.82	4
M109A2 (SPH) (USA)	3.28	0,61	5.4
M109A6 Paladin (SPH) (USA)	3.28	0,61	5.4
AS-90 (SPH) (Britain)	3.00	0,62	4.8
PzH 2000 (SPH) (Germany)	3.50	0.70	5

Another possible recognition signature can be the ratio of the turret length b to the overall vehicle length l.

In a SPH, the turret is long and tall due to the large dimensions of the artillery installation, extending all the way to the rear end of the hull. In a tank, on the other hand, to ensure high maneuverability, low visibility, and enhanced crew protection – with the crew occupying a significant portion of the turret – the turret is relatively flat and low compared to the overall height of the tank. Its length b, measured along the longer side of the hull, is significantly smaller than the hull length, so that the rear edge of the turret is clearly set back from the rear end of the hull by a noticeable distance. This distance δ , measured from the rear edge of the turret to the rear of the hull, can also serve as a recognition signature distinguishing a tank from a SPH.

4.1.2. Probability of recognition using a single signature dimension. Given that an image has a finite size resolution, the probability of recognition by shape depends not only on the signature size but also on the limiting size that can still be resolved by the computer. Specifically, the size of the characteristic detail must exceed the threshold resolution size. In this case, the probability of recognizing the target as an object of interest in an electronic image is a geometric Bayesian probability, provided that the size of the characteristic detail of the vehicle in the image is larger than the threshold resolution size of the image.

It is important to note that the term size of an image element may refer either to its two-dimensional (2D) size – that is, the area s of the image element – or to its linear, one-dimensional (1D) dimensions. The concept of *dimensionality* (2D or 1D) of an image element's size is a key distinction in defining the size-based conspicuity indices for target *detection* and target *recognition* in an image.

When accomplishing the visual task of detection during reconnaissance we deal with 2D dimensions, such as the image area S, the area of the target s as an object of interest in the image, and the threshold area of the lower limit of resolution s_{th}.

When accomplishing *the visual task of target recognition,* we deal with comparing 1D dimensions that define the *shape* of the combat vehicle or the *anisometry of the forms* of its structural elements.

In the case of an image obtained from ground-based cameras, during recognition we deal with pairs of 1D dimensions of image elements (length, diameter, thickness, distance) such as

• length l and height h of the vehicle;

• length l and diameter (thickness) d of the vehicle's gun barrel;

• turret length b and hull length l;

• distance from the rear edge of the turret to the hull's rear end, and so on.

If the image is obtained from UAV cameras or from the frontal projection of the vehicle, then, in addition to these 1D dimensions, the width w of the image element must also be considered.

When recognizing by 1D dimensions, the image-characterizing lengths are the length of the SEZ L and the threshold length of the lower resolution limit l_{th}. It is important to note that, unlike the detection task, in recognition L is not the length of the entire image frame but the length of the SEZ – the zone in the image where the given recognition signature is expected. The reason is that once a combat vehicle has been detected in the image, the observer no longer needs to scan their gaze across the whole target expectation zone (TEZ) along the entire line parallel to the horizontal side of the image frame, as was necessary for detection. For vehicle recognition, it is sufficient to scan the gaze within the signature expectation zone (SEZ), which covers all extreme points of the target contour – for example, a rectangle with sides parallel to the image frame that encloses the contour points so that the most distant of them lie on its sides. In fact, during recognition the target area $s=l\,h$ becomes the SEZ area, and its length and height become the horizontal and vertical 1D dimensions of the SEZ if the given signature is observed against the vehicle hull. However, in the case when the signature is located significantly outside the vehicle contour – for example, the gun barrel of an SPH in Figure 2 – the SEZ is defined somewhat differently. Such an example will be analyzed below. For example, if the recognition of a combat vehicle is performed by the length of the gun barrel, then the probability of recognition is determined as:

$$P_l = \begin{cases} \dfrac{1-\dfrac{l_{th}}{l_{sg}}}{1+\dfrac{l_{th}}{l_{sg}}\left(1-2\dfrac{l_{sg}}{L}\right)} & at \;\; l_{th} \le l_{sg} \le L;\; d \ge l_{th} \\ 0 \quad at \quad 0 \le l_{sg} \le l_{th};\; d \le l_{th} \end{cases}, \qquad (22)$$

where l_{sg} is the signature length, the gun barrel in this case; d is its diameter; L is the length of the horizontal side of the SEZ rectangle; and l_{th} is the threshold (minimum) length (per Nyquist) of an image element that is still visualized on the display.

For the image shown in Figure 2a, the SEZ covers almost the entire frame. According to formula (17), the image frame length is $L=(7.75/2.54)\times 96=293\,px$, while the SEZ length, corresponding to the length of the SPH in the image, is l=254px. In Figure 2a, the gun barrel of the SPH is clearly visible, and there is no need to enlarge the image using the “zoom” operation. The measured horizontal projection of the SPH gun barrel length in Figure 2a is l_{sg} = 169px. By substituting l = 254, l_{sg} = 169 and l_{th} = 2 into formula (22), we obtain the size-based conspicuity index of the SPH gun barrel in the image (Fig. 2a) $P_l=0.996$. Intuitively, such a high value of the size-based conspicuity index (i.e., recognition probability) indicates an unambiguous recognition of the SPH as a combat armored vehicle.

Expression (22) takes into account that the key conspicuity feature for recognizing a tank is the length of its gun barrel, but at the same time the barrel diameter d in the image must be greater than the linear resolution $d > l_{th}$ of the monitor. If $d \le l_{th}$, then according to expression (22) the probability of detection is zero $P_l\big|_{d\le l_{th},\,\forall l}=0$ for any barrel length l_{sg}.

4.1.3. Probability of recognition using the sizes of two signatures. An important distinction of the *recognition* procedure from the *detection* is that in the detection the target is perceived as a whole. That is, the variable that characterizes the target size is its area. In recognition we deal with the anisometry of the combat vehicle's shape, i.e., with at least a pair of linear dimensions such as: the tank's length and height, the length and diameter of the combat vehicle's gun barrel, the vehicle length and turret length, turret height and hull height, etc. Each of the two lengths l_1 and l_2 belonging to the same pair is characterized by its own detection probability P_{l1} and P_{l2}, respectively. For the target to be recognized with probability $P_R \neq 0$, it is necessary and sufficient that both recognition events by the two size features in both directions occur simultaneously with nonzero probabilities. If at least one of the two probabilities $P_{l1} = 0$ or $P_{l2} = 0$, or both together, equals zero $P_{l1} = P_{l2} = 0$, then we have $P_R = 0$ that is:

$$P_R\big|_{P_{l_1}=0} = P_R\big|_{P_{l_2}=0} = P_R\big|_{P_{l_1}=0,\, P_{l_2}=0} = 0\,. \quad (23)$$

Condition (23) means that the events with probabilities P_{l1} and P_{l2} are multiplicative with respect to each other and to the resulting event P_R therefore

$$P_R = P_{l_1} \cdot P_{l_2}\,. \quad (24)$$

To within the physical meaning of the variables, the quantities P_{l1} and P_{l2} are determined by expression (22), that is

$$P_{l1,2} = \begin{cases} \dfrac{1-\dfrac{l_{th}}{l_{1,2}}}{1+\dfrac{l_{th}}{l_{1,2}}\left(1-2\dfrac{l_{1,2}}{L_{1,2}}\right)} & \forall l_{th} \le l_1 \le L_1;\ l_{th} \le l_2 \le L_2 \\ 0 \quad \forall \quad l_1 \le l_{th};\ l_2 \le l_{th} \end{cases}, \quad (25)$$

where L_1 and L_2 are the lengths of the sides of the SEZ rectangle, i.e., the lines in the image along which the gaze scanning is performed.

By substituting expressions (25) into formula (24), we obtain

$$P_R = \begin{cases} \dfrac{(l_1-l_{th})(l_2-l_{th})}{\left[l_1\left(1-2\dfrac{l_{th}}{L_1}\right)+l_{th}\right]\left[l_2\left(1-2\dfrac{l_{th}}{L_2}\right)+l_{th}\right]} & \forall l_{1,2} \ge l_{th} \\ 0 & \forall l_{1,2} < l_{th}, \end{cases} \quad (26)$$

After transformation, expression (26) for $\forall l_{1,2} \ge l_{th}$ can be rewritten in the form

$$P_R = \frac{1-\dfrac{s_{th}}{s_{12}}\left(\dfrac{l_1+l_2}{l_{th}}-1\right)}{1+\dfrac{s_{th}}{s_{12}}\left[1+\dfrac{l_1+l_2}{l_{th}}+2\dfrac{s_{12}}{S}\left(2-\dfrac{L_1+L_2}{l_{th}}-\dfrac{L_1}{l_1}-\dfrac{L_2}{l_2}\right)\right]}, \quad (27)$$

where $s_{th} = l_{th}^2$ – is the threshold area of resolution, and $s_{12} = l_1 l_2$ is the area of the signature (the area of the gun barrel of a tank or SPH, for example) by which the combat vehicle is recognized against the SEZ of area $S = L_1 L_2$.

Comparison of expressions (16) and (27) shows that the probability of target *detection* P_s by its size feature and the probability of its *recognition* P_R by the size feature of a particular signature in a digital image differ by at least the following characteristics.

First, the area of the recognition signature in expression (27) is smaller than the target area S in expression (16), which by itself reduces the probability of recognition compared to the probability of detection.

Second, since the CAV as the object of interest (target) in the image has already been detected, the search for the signature is carried out not within the TEZ but within a much smaller the SEZ bounded by the vehicle's contours. This, in turn, increases the probability of detecting the signature – that is, the probability of recognizing the vehicle by that signature.

Third, expression (27) includes not only the signature area s_{12} but also its linear dimensions l_1 and l_2, which implicitly accounts for the anisometry of the signature's shape.

Let us determine the probability of detecting the barrel in image Figure 2a, from which we find $l_1 = 169\,px$, $l_2 = 11\,px$. The length and height of the SEZ rectangle are estimated to be half of the vehicle's length and height together with the barrel, whence $L_1 = 130\,px$, $L_2 = 80\,px$ and for $l_{th} = 2\,px$ substituting these data into expression (26), we obtain $P_R \approx 0.71$.

There is no doubt that, at a probability $P_l = 0.71$ the detection of the signature (the SPH barrel in Fig. 2a, in this case) by the size feature will be successful. However, the sizes of the vehicle and its characteristic details in the image are not always so large relative to the image frame dimensions and the threshold resolution l_{th}. Therefore, the probability of solving the visual recognition task can take much smaller values for significantly smaller l_1 and l_2. Hence, similar to the criterion for successful accomplishment of the *detection task*, the question arises of a criterion for successful accomplishment of the target *recognition task* by the size feature of a signature on an electronic image.

4.2. Criterion for successful target recognition by the size feature of a signature.

Similar to the criterion for successful accomplishment of the target detection task by the size feature on an electronic image, we introduce a criterion for successful accomplishment of the target *recognition task* by the size feature of its signature on the same image

$$P_R > \left(P_R\right)_{50} = \frac{1}{2}. \qquad (28)$$

To determine the signature dimensions $(l_1)_{50}$ and $(l_2)_5$ that satisfy the condition $(P_R)_{50} = 1/2$, it is necessary to solve the equation

$$\frac{\left((l_1)_{50} - l_{th}\right)\left((l_2)_{50} - l_{th}\right)}{\left[(l_1)_{50}\left(1-2\frac{l_{th}}{L_1}\right)+l_{th}\right]\left[(l_2)_{50}\left(1-2\frac{l_{th}}{L_2}\right)+l_{th}\right]} = \frac{1}{2}. \qquad (29)$$

$\forall\, l_{1,2} \geq l_{th}$

Equation (29) contains two unknown variables, $(l_1)_{50}$ and $(l_2)_{50}$. However, for a specific type of CAV to be recognized by a given signature, the ratio $k = l_1/l_2$ is generally known. Hereinafter (without loss of generality), for convenience, we assume that l_1 is the longer and l_2 is the shorter dimension of the signature, i.e., $l_1>l_2$. For the gun barrel of an SPH or tank, k is simply the number of calibers corresponding to its length, the tabulated value of the barrel length measured in its calibers. Therefore, substituting $(l_1)_{50} = k(l_2)_{50}$ into equation (29), we obtain an equation with a single unknown $(l_2)_{50}$. Equation (29) thus reduces to a quadratic equation

$$a\left((l_2)_{50}\right)^2 - b(l_2)_{50} - c = 0, \qquad (30)$$

where
$$\begin{cases} a = k\left[2-\left(1-2\dfrac{l_{th}}{L_1}\right)\left(1-2\dfrac{l_{th}}{L_2}\right)\right], \\ b = 2\,l_{th}\left[k\left(1-\dfrac{l_{th}}{L_1}\right)+\left(1-\dfrac{l_{th}}{L_2}\right)\right], \\ c = 3\,l_{th}^2, \end{cases}$$

which accordingly has two roots. Since the free term in this quadratic equation is negative, by Vieta's theorem we conclude that one of the roots is positive and the other is negative. Because $(l_2)_{50}$ has the physical meaning of a length, its value cannot be negative. Therefore, among the two roots of the quadratic equation, the positive one is taken. The analytical expression for $(l_2)_{50}$ is rather cumbersome and is not presented here. Instead, expanding it into a Taylor series with respect to the small parameter l_{th} and retaining only the first-order term, we obtain

$$(l_2)_{50} \approx \frac{3}{2}\left(1+\frac{1}{k}+\sqrt{1+\frac{14}{k}+\frac{9}{k^2}}\right)l_{th}. \qquad (31)$$

From expression (31), one can determine the range within which the quantity $(l_2)_{50}$ varies for $k \in [1;\infty]$. For the value $k = 1$, which corresponds to an isometric signature in the form of a circle or a regular polygon (triangle, square, pentagon, n-gon), we obtain $\left.(l_2)_{50}\right|_{k=1} \approx 12 l_{th}$. For a markedly elongated signature element $k \gg 1$, we obtain $\left.(l_2)_{50}\right|_{k=1} \approx 6 l_{th}$. Thus, we find that the criterion for the successful solution of the recognition task varies within $(l_2)_{50} = [6;12]l_{th}$.

For the target recognition task by the sizes of its signatures to be successfully accomplished, the following condition must be satisfied

$$l_2 > (l_2)_{50}, \qquad (32)$$

that is, the smaller of the signature dimensions in the image must exceed the threshold signature size that ensures a 50% probability of its detection. Expression (32) is the analytical form of the criterion for successful target recognition by the sizes of its signatures.

To compare the values of the success criteria for *detection* and *recognition* tasks, we introduce the notation $(l_2)_{50} = l_{50}^R$, where the superscript R (from recognition) denotes recognition.

Let us recall that the criterion for successful accomplishment of the target *detection task* is given by expression (20) through the area s_{50}. For convenience of comparison, we introduce the notation $l_{50}^D = \sqrt{s_{50}}$, where the superscript D (from detection) denotes detection. From expression (20), we obtain

$$l_{50}^D = \sqrt{3} \times l_{th}. \qquad (33)$$

Since according to expression (31) the value of $l_{50}^R(k)$ is a function of the parameter k, to compare it with the corresponding parameter l_{50}^D for the detection task, we perform averaging of the quantity $l_{50}^R(k)$ over the interval of values $k \in [k_1;k_2]$. Taking into account expression (31), the averaged value is determined as

$$\overline{l}_{50}^R = \frac{3}{2}\frac{l_{th}}{k_2-k_1}\int_{k_1}^{k_2}\left(1+\frac{1}{k}+\sqrt{1+\frac{14}{k}+\frac{9}{k^2}}\right)dk. \qquad (34)$$

For the given values of $k \in [1;40]$, we obtain

$$\overline{l}_{50}^R \approx 7 \times l_{th}. \qquad (35)$$

And thus, from expressions (33) and (35), we obtain the comparative estimate

$$l_{50}^D : \overline{l}_{50}^R \approx 1:4, \qquad (36)$$

which correlates well with the empirically established relationship (5) for the Johnson and TTP criteria of successful accomplishment of visual detection and recognition tasks

$$x_{50}^{D} : x_{50}^{R} \approx 1:4 . \quad (37)$$

The quantitative agreement between the success criteria for solving the target detection and recognition tasks by the size feature, given by expression (36) and theoretically derived in this study, and the corresponding empirically obtained criteria given by expression (37), confirms the adequacy of the developed approach to target recognition based on the geometric Bayesian probability of detecting a specific recognition signature by two dimensions.

4.3. The role of the proposed method for estimating the probability of target recognition by the sizes of its signatures in a digital image in target recognition performed by a human operator and by artificial intelligence (AI).

It is important to note that the method developed in this study for estimating the probability of target recognition by the sizes of its signatures in a digital image can serve both as a means of assessing the adequacy of AI-based target recognition and as an independent alternative tool for target recognition by a human operator during reconnaissance.

4.3.1. A posteriori verification of recognition adequacy. Regarding the first application – the verification of target recognition adequacy using AI – it should be noted that the initial iteration of AI-based recognition, for example of armored vehicle samples in an image, often includes objects marked by the AI as targets that are, in fact, not targets. In such cases, the operator must manually filter out falsely marked objects (e.g., trucks or buildings instead of armored vehicle samples).

The causes of erroneous object recognition as targets by AI may vary. One of them is the neglect of the fact that not every object in an image can be successfully recognized. In the previous subsection, it was shown that for the recognition task based on the sizes of target signatures to be considered successfully accomplished, the condition $l_2 > (l_2)_{50}$, must be satisfied – that is, the smaller of the signature dimensions must exceed the threshold signature size that ensures a 50% probability of detecting that signature. In other words, to avoid recognition errors associated with insufficient signature sizes, it is necessary to ensure that condition (32) is satisfied for the recognition signatures. However, to do so, the AI must first detect the signatures on the electronic image of the object of interest by which it can be recognized, for example, as an armored vehicle sample. This task may be too ambiguous and hard for AI. Nevertheless, this is not required. To verify whether condition (32) is met, one can use the known ratio between the size of a signature and the overall size of the target, which is typically a tabulated value. For instance, by using tabulated ratios of a tank turret height to its total height, the AI can apply the recognition success criterion not to the size of specific signatures of an armored vehicle sample but to its overall dimensions – such as height or width – in the digital image.

Below, in the following paragraphs, this possibility will be illustrated with specific examples. However, for the reader’s convenience, it is first necessary to briefly explain the relationship between the number of pixels in an image and the metric units of size (length, height, width in millimeters or centimeters). To convert the number of pixels into metric units (e.g., millimeters), in the examples given above, we used the image resolution value in dpi (dots per inch), which is commonly used to describe the quality of printed images (on paper) and is included in the image file metadata. If the image is viewed on a monitor, instead of dpi, one should use the monitor resolution value in ppi (pixels per inch) – the number of pixels per inch. Some programs display images on a monitor while taking the image’s DPI into account, while others do not. For example, Microsoft Word considers the image’s DPI when inserting it into a Word document: inserting an image of 300×300 pixels with DPI=300 will display it as 1×1 inch (2.54×2.54*cm*). Knowing the screen width and height in pixels and its diagonal in inches, one can calculate the pixel size in metric units, such as in millimeters.

4.3.2. Examples of adequacy assessment and recognition of an armored vehicle sample by its signatures. For example, if an armored vehicle sample (a tank or an SPH) is recognized by the presence of a turret, then according to expressions (32) and (35), the turret height in the image must exceed $\bar{l}_{50}^{R} \approx 7 \times 2 = 14\left(px\right)$. According to expression (17), for an image with a resolution of 96 *dpi*, 14 pixels correspond to 14×2.54/96≈3*mm*. Considering that for an armored vehicle sample the ratio of its total height to the turret height is approximately 3:1, it follows that, in order for the armored vehicle to be recognized by the presence of a turret, its image must have a total height of at least 14×3 = 42*px*, which on an image with a resolution of 96 dpi, according to expression (17), corresponds to 42×2.54/96≈10*mm*. Given that for tanks and SPHs a typical ratio of hull length to vehicle height is about $l/h \approx 3$, it follows that, when viewed from the side (as in Fig. 2c), an image of an armored vehicle sample with a turret height of 3*mm* and a total height of 10mm on an image with a resolution of 96*dpi* would correspond to a hull length of approximately 30*mm*.

If the armored vehicle sample is observed not from the side but in a frontal projection, as shown for example in Figure 3, it can be recognized by using pairwise ratios, such as the ratio of the turret height to the total vehicle height, the turret height to the hull height (from the ground to the lower level of the turret), and the total vehicle height to its width.

a b c d

Fig. 3. **Image of the T-64BV tank captured by a digital camera with 1× magnification at distances of a – 10 m; b – 20 m; c – 50 m; d – 100 m**

In the first two cases, recognition is performed based on the sizes of the recognition signatures (the heights of the turret, hull, and total height), while in the third case – based on the characteristic feature of the vehicle's shape anisometry (height and width). If the recognition of an armored vehicle sample, such as the T-64BV tank, is carried out using its image in a frontal projection, it can be considered that the ratio of the turret height to the total height typically equals 1:3, and the ratio of the turret height to the hull height equals – 1:2. Then, for successful recognition by the size features of the signatures, the turret height in the image should be at least 14*px* (≈3*mm* for an image with a *resolution of* 96*dpi*), while the hull height and total vehicle height would then be 14×3=42*px* (≈9*mm*) and 14×2=28*px* (≈6mm), respectively.

For the T-64BV tank, the ratio of its width to height measured from Figure 3 is *w/h*=1/6, which agrees well with the tabulated value *w/h*=3460*mm*/ 2190*mm*=1.58 [26]. Therefore, for *successful recognition* of the tank by the characteristic feature of its shape anisometry in the image, for a tank image with a total height of 14×3=42*px* (≈9*mm*), the vehicle width should be at least 68*px* (≈15*mm*). The estimates of minimum signature sizes and shape anisometry of AFV (including armored vehicles) samples in the image, according to the recognition success criterion (32), should be applied not only for recognition by a human operator when viewing the image on a monitor but even more so for AI-based recognition. To achieve this, using the tabulated ratios $l_1 = kl_2$ between the signature dimensions l_2 of an AFV sample and its overall dimensions l_1 (height *h* or width/ length *l* of the vehicle), the criterion (32) can be reformulated as:

$$l_1 > (l_1)_{50}. \quad (38)$$

The AI-based algorithm for recognizing armored vehicle objects in electronic images should reject image elements whose dimensions do not satisfy criterion (38), as such elements do not ensure a recognition probability higher than 50%.

In addition, it should be taken into account that recognition errors (especially in AI-based recognition) may result not only from insufficient sizes of the recognized object in the image but also from excessively low values of its luminance contrast. In other words, the second necessary condition for successful recognition is condition (13). The AI algorithm should therefore reject image elements whose recognition signature contrast does not satisfy criterion (13). For convenience and to simplify the evaluation procedure, instead of measuring the luminance contrast of individual signatures, the overall luminance contrast of the armored vehicle sample can be used. Thus, the AI algorithm should reject recognition objects whose contrast is lower than $|K_{50}| = 0.14$ (for $|K_{th}| = 0.05$), as such objects do not ensure a recognition probability by the luminance contrast feature higher than 50%. It is worth noting that in some publications the threshold contrast value is considered to be more than twice lower, $|K_{th}| = 0.02$, for which, according to (12), one finds $|K_{50}| = 0.05$. Which of the two values of $|K_{th}|$ should be used in the estimation of $|K_{50}|$ deserves additional studies. For the time being, the higher value $|K_{th}| = 0.05$ is preferable to ensure correct recognition.

The procedure for experimental determination of the detection and recognition probabilities of an armored vehicle sample will be illustrated in the following (fifth) section of the paper, using as an example images of the main battle tank (MBT) T-64BV, evaluated by the sizes of its signatures on images obtained at different distances from the target.

5. Experimental determination of detection and recognition probabilities of an MBT sample based on the sizes of its signatures as a function of distance

To determine the detection and recognition probabilities of the MBT sample (T-64BV) based on the sizes of its signatures as a function of distance, an experiment was conducted with the aim of collecting images of the T-64BV tank at distances ranging from 10*m* to 750*m*. Examples of some of the obtained images, specifically for distances D=10, 20, 50, 100m, are shown in Figures 3a-d, respectively.

5.1. Experimental conditions.

5.1.1. Image acquisition was carried out using a Nikon Coolpix L830 Black digital camera at the International Peacekeeping and Security Center training range during winter, under overcast daylight conditions. Images were obtained over the distance range of [0;100]m with a 10m step, over [100;500]m with a 50m step, and over [500;750]m with a 100m step; the final image was captured at a distance of 750m.

5.1.2. Measurement of signature dimensions of the armored vehicle sample – the tank's height (h), width (w), turret height (h_t), and hull height (h_h) – on the collected images (in pixel units) was carried out using the QTDDI computer application [27] and can also be performed in commonly available programs, for example, in XnView.

The probability of detection was determined using expression (16), where $s = h \times w$ is the area of the rectangle enclosing the target, with width w and height h $s_{th} = 4px^2$ is the threshold 2D resolution of the monitor according to Nyquist, and $S = L \times H$ is the image size in pixels. An advantage of a digital image over an analog one is that it can be magnified on the monitor using the zoom operation, which significantly increases the accuracy of measuring signature dimensions. The absolute measurement error of signature dimensions can be reduced to 1 pixel by using a brightness profile graph plotted along a horizontal line through the target and the pixel brightness table for the given image, in accordance with the QTR-EI index determination methodology [6].

5.2. Results of measurements and calculations.

5.2.1. Probability of detection. The probability of detection of the tank was calculated using formula (16), where s=w×h – the area of the tank on the image, and S=L×H=4608×3456(px^2) is the total image area obtained at a distance D from the tank. Since the ratio between the tank's width w and height h is a constant, it was unnecessary to measure both parameters on all images. Measuring only one – preferably the more distinguishable one – is sufficient, while the other can be computed using the known proportion w/h=1.6. This approach not only simplifies and accelerates data acquisition but also improves computational accuracy. As the distance D increases, both the apparent size and contrast of the tank's image decrease, thereby reducing measurement precision. Among the two dimensions, width w exhibits higher contrast and is therefore more reliable for measurement. In the collected images (some shown in Fig. 3), width w was measured directly, while the height was computed as h=w/1.6. To demonstrate the validity of this simplification, experimentally measured height values h(D) is shown as red open circles in Figure 4, while black squares represent the corresponding computed values obtained from width data using h=w/1.6.

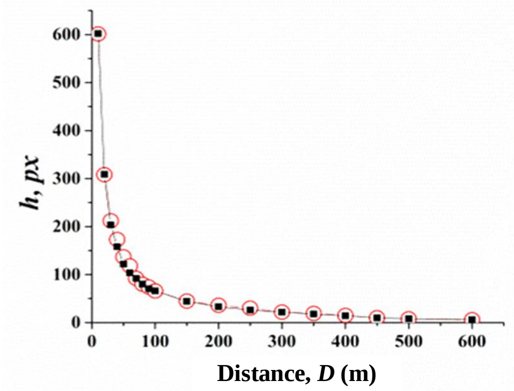


Fig. 4. **Experimentally measured values (red open circles) of the tank height h and the corresponding values calculated as h=w/1.6**

The coincidence of the experimentally measured and calculated values shown in Figure 4 confirms the validity of the simplified measurement procedure.

The dependence s(D), expressed in square pixels, was computed from the experimentally measured values of w and h (Fig. 5). For convenience, the vertical axis is presented in a logarithmic scale.

Similar to Figure 4, in Figure 5 the red open circles represent the values of the tank image area obtained from independently measured w and h, while the black squares correspond to the values derived from width data using the expression $s = w^2/1.6$.

The agreement between both datasets in Figure 5 once again confirms the adequacy of the simplified measurement procedure. Therefore, this approach will be further used to obtain the dimensions of other signatures.

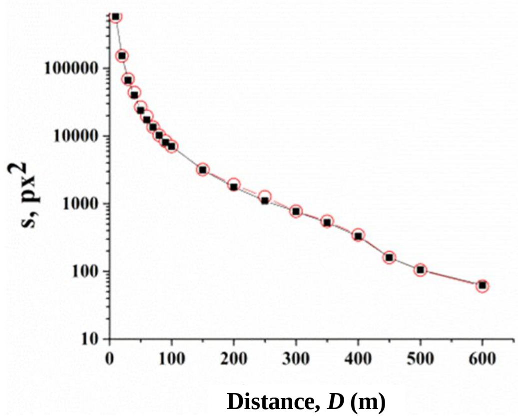


Fig. 5. **Experimentally obtained dependence (in logarithmic scale) of the tank image area s(D) on the distance D from the tank: red open circles – values obtained as s=w×h from independently measured w and h; black squares – corresponding values calculated as $s = w^2/1.6$**

The probability of tank detection in images acquired at various distances D, calculated using equation (16) from the data presented in Figure 5, is shown in Figure 6. It can be seen from the figure that the probability of detecting the tank by the distinctive feature of its size remains higher than 0.5 (50%) even at a distance of 600m.

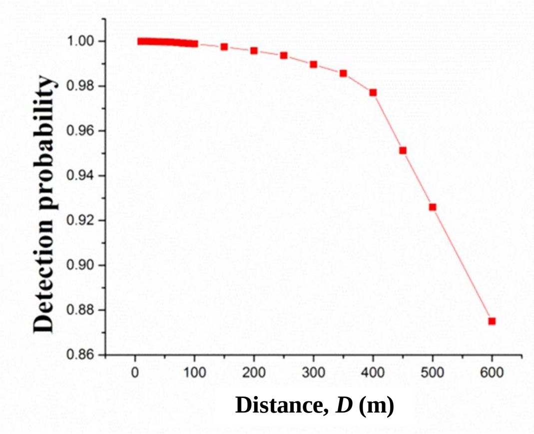


Fig. 6. **Probability of detecting the T-64BV tank in an image captured with a digital camera at a zoom factor of ×1 as a function of the distance to the tank**

Analyzing the images corresponding to different distances to the tank, it can be concluded that with increasing distance not only does the apparent size of the tank on the image decrease, but also its contrast, i.e., the target visibility index V.

Furthermore, as the distance increases, the target conspicuity index on a cluttered background $P_{clutter}$ decreases. Due to the significant reduction in contrast with distance and the diminished conspicuity of the target against a cluttered background, it becomes impossible to detect the tank in images acquired at distances larger than 600m, even though the probability of detection by size remains above 0.5 (50%).

The reason lies in the fact that, according to equation (8), the target saliency index Γ on an electronic image is the product of the target visibility index (conspicuity by brightness contrast) V, the size conspicuity index, and the background clutter conspicuity index $P_{clutter}$. Thus, because of the low overall saliency resulting from small values of V and $P_{clutter}$, the target's visibility on images taken at distances D>600m is too low for successful detection – therefore, points corresponding to D>600m are absent on the graph (Fig. 4).

To illustrate that the low saliency of the target at distances D>600m is not a consequence of its size conspicuity, we constructed model images of the tank in Mathematica as rectangular shapes with a visibility index V=1 on a uniform white background, thereby ensuring $P_{clutter}$. The model rectangles had the same dimensions (in pixels) as those shown in Figures 4 and 5, which were used to calculate the detection probabilities presented in Figure 6. Accordingly, for this model case, in accordance with equation (8), the target saliency index reduces to the size conspicuity index $\Gamma\big|_{V=1,\,P_{clutter}=1} = P_S$. Model images of the target with sizes corresponding to the tank's dimensions at distances within the range [10;600]m is shown in Figure 7. It is evident from Figure 7 that even the smallest rectangle with dimensions $w \times h = 10 \times 6(px^2)$, corresponding to a distance $D = 600m$, remains clearly visible – i.e., it can still be detected by the distinctive feature of its size on the image – consistent with the data presented in Figure 6.

5.2.2. Probability of recognition. The probability of recognition of the tank was calculated using expression (26), based on the dimensions of two signatures: the height of the turret h_t and the height of the hull h_h. Since at the recognition stage the target has already been detected, the SEZ is simply a rectangle with sides equal to the tank's width w and height h. When searching for the turret and hull on the image, the visual scanning is performed along a vertical line from the ground to the top point of the turret (or vice versa); therefore, in equation (26), we take $L_1 = L_2 = h$.

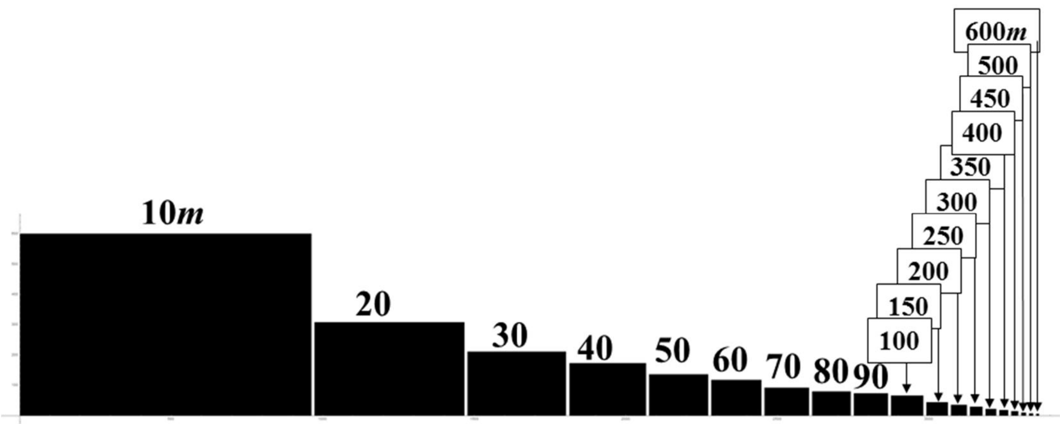


Fig. 7. **Model images of the tank represented as rectangles with visibility indices V=1, $P_{clutter}$=1, and dimensions corresponding to those used for calculating the detection probability at the distances (in meters) indicated above each image**

Dependencies $h(D)$, $h_t(D)$, and $h_h(D)$ – corresponding to the total height of the tank, the turret height, and the hull height on the images as a function of distance D – are shown in Figure 8.

Since the ratios between the overall height h, the turret height h_t, and the hull height h_h of the tank are constant tabulated values $h_t/h = 1/3$, $h_h/h = 2/3$, there is no need to perform independent measurements of h_t and h_h. These parameters can be derived from previously obtained height values h, which in turn can be calculated from the measured tank width using $h=w/1.6$.

The data presented in Figure 8 were used to calculate the probability of recognition of the tank, shown in Figure 9 with blue dots, based on the size features of the turret and hull as recognition signatures. Additionally, in Figure 9, the probability of recognition by the gun barrel length and diameter is plotted with red triangles. For comparison, the probability of detection data from Figure 6 are included with green squares.

From Figure 9, it is evident that, as expected, at short distances (~10*m*) the probabilities of detection and recognition are close in value and nearly equal to 1. However, as the distance increases, the recognition probability becomes significantly lower than the detection probability.

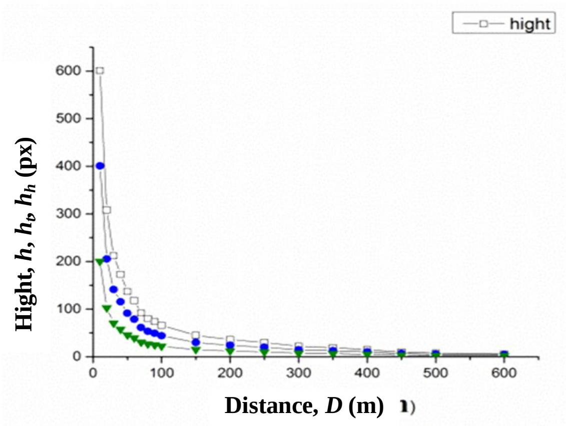


Fig. 8. **Dependencies h(D) – open squares; $h_t(D)$ – blue filled circles; $h_h(D)$ – green filled triangles**

At distances $D \geq 350m$, the recognition probability based on the turret and hull heights drops below 0.5 (50%), indicating that recognition becomes unreliable or even impossible, even at high zoom magnifications.

This observation is confirmed in Figure 10, where the bottom row shows 500% magnified image fragments taken at distances of 300*m* and 400*m*. Indeed, at D=300m the turret and hull of the tank can still be distinguished separately, whereas at D=400*m* the tank can no longer be reliably recognized by any of its signatures.

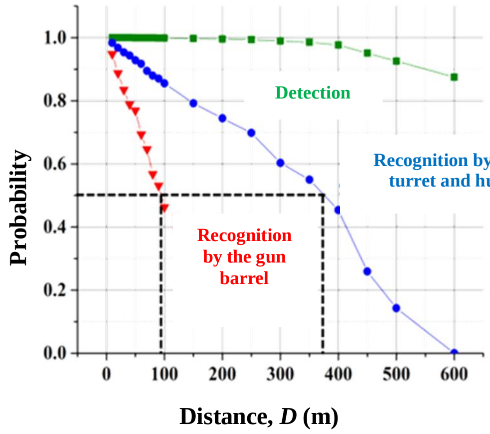


Fig. 9. **Probabilities of recognition based on turret and hull height (blue dots), recognition based on gun barrel dimensions (red triangles), and detection based on image area (green squares) as functions of distance to the tank**

This conclusion indicates an agreement between the recognition range limit (~350*m*) corresponding to the theoretically calculated 50% recognition probability (Fig. 9) and the empirically observed recognition limit (300-400*m*) established by visual inspection of the photographs.

Another signature by which the tank can be recognized is its gun barrel. The probability of barrel detection, calculated using expression (26) from its length l_1 and diameter l_2 on images obtained at various distances, is shown in Figure 9 as red triangles.

In the calculations, considering that the tank is already detected in the image and the visual scanning is performed not across the entire frame but only within the SEZ, the SEZ is taken as a rectangle in which the tank gun axis lies along its diagonal. From Figure 9, it is evident that the recognition probability based on the gun barrel signatures drops below 50% at distances D>100*m*.

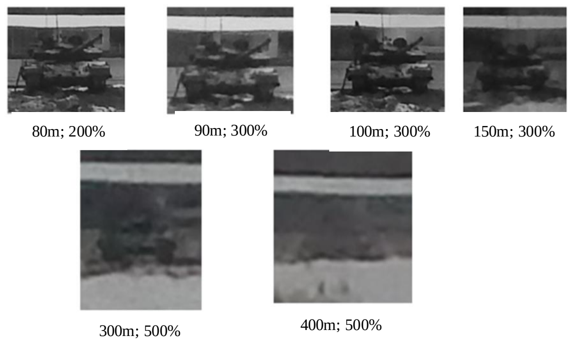


Fig. 10. **Fragments of magnified images at different distances. The magnification factor and the corresponding shooting distances are indicated below each image**

This conclusion is consistent with the visual inspection of the magnified images. Indeed, in the image taken at a distance of 100m, the gun barrel is still visible, whereas at 150m it becomes difficult to distinguish it with confidence. Thus, it can be concluded that the theoretically derived criterion of recognition success is consistent with the empirical visual observations of the photographs, also with respect to the signature features of the tank barrel.

6. Research prospects

It should be noted that this study focuses on the probability of recognition of military equipment samples based on the dimensions of their signatures. The influence of the luminous contrast of signatures on the probability of target recognition will be the subject of our subsequent publications. At this stage, it can be stated that the overall probability of recognition based on both the size and contrast of the signatures is the product of the respective probabilities. Therefore, for successful recognition, both conditions (32) and (13) must be simultaneously satisfied – corresponding to the signature dimensions and their luminous contrast, respectively.

In this work, a methodology for determining the probability of target recognition based on the geometric dimensions of its signatures has been developed, and criteria for the successful accomplishing visual detection and recognition tasks based on distinctive features have been established. The method was applied to images obtained in the visible spectrum using a digital camera with a single zoom factor. In addition, during the experiment described above, we obtained a set of images using the same camera with maximum zoom, as well as with a mobile phone camera at both 1× zoom and maximum zoom. Comparative studies of the detection and recognition probabilities for images obtained under these different optical conditions are ongoing and will form the basis of our future publications.

An independent interest is the application of the theoretical approach developed in this work to images obtained with thermal imagers, including at different zoom levels. Corresponding thermal images were also acquired during our experiment and are currently being processed.

Another promising possibility afforded by the digital format of images is the ability to measure the distance to a target from its size in the image, measured in pixels using the image brightness table. Figures 4, 5 and 7 demonstrate a clear one-to-one functional relationship between the target size l (in pixels) in a digital image and the distance D to it. For a given sightseeing complex of an AFV sample, this functional dependence (D) can be tabulated experimentally or obtained theoretically in the form of an analytical relation. Using the brightness-profile plot along a line drawn through the target, the size l of the target on the image can be measured with an accuracy up to 1 px. Thus, having the target size l on the electronic image and the dependence $l(D)$, the distance to the target can be determined.

Conclusions

The capability to obtain information about the target-background situation in the form of digital images opens previously unavailable opportunities for real-time battlefield monitoring. However, the enormously bulky stream of video content acquired from various sources requires equally enormous human effort and computational resources to analyses these datasets. It is no secret that, in most cases, the lion’s share of the collected video content contains no information about a target or is of insufficient quality. The development of theoretical basis for assessing the quality of target data acquisition (detection, recognition and identification) based on quantitative measurements of target displaying in digital images enables automated selection of the highest-quality digital images without operator involvement.

In this study, we analyzed existing approaches to measuring the probability of target detection on both analog and electronic images. As a continuation of our previously developed the QTDDI approach aimed at target detection, this work extends the concept to the next stage of the target data acquisition process – recognition. To achieve this, we theoretically derived an analytical dependence describing target recognition probability based on the dimensions of its signatures in a digital image. The derived analytical expressions for the probability of detection (16) and probability of recognition (26) serve as the theoretical basis for a method of evaluating detection and recognition probabilities according to the dimensional features of targets in digital images (including those obtained from multi-channel sightseeing systems of armored vehicles). Based on the proposed method, we developed success criteria for accomplishing visual tasks of target detection and recognition, which can be implemented using brightness data tables of digital images.

The application of these criteria to specific images enables automated filtering of low-quality digital imagery at the input stage, before initiating the process of AI-based target search and recognition. This approach can significantly reduce image-processing time, improve recognition efficiency, and minimize false detections.

To verify the theoretical results obtained, we conducted an experimental study involving the collection of digital images of the T-64BV main battle tank at various distances using different imaging devices. For the images acquired with a digital camera at a 1× zoom level, we calculated – from the quantitative data of the image brightness tables – the probability of target detection based on its size and the probability of recognition based on the dimensions of its signatures as functions of distance to the target. The agreement between the 50% success criterion (for detection and recognition probability) and the visual analysis of the experimental images confirms the adequacy and reliability of the proposed approach.

References

1. Khaustov D.Ye., Kyrychuk O.A., Stakh T.M., Khaustov Ya.Ye., Burashnikov O.O., Ryzhov Ye., Vlokh R. and Nastishin Yu. A. (2023), Complex-scalar and complex-vector approaches for express target-oriented image fusion. *Ukrainian Journal of Physical Optics*. Lviv. Vol. 24. 1. pp. 62-82. DOI: https://doi.org/ 10.3116/ 16091833/24 /1/62/2023

2. Khaustov D. Ye., Nastyshyn Yu. A., Khaustov Ya Ye., Malynych S. Z. and Hryvachevskyi A. P. (2021), "Formuvannia ta obrobka zobrazhen u prytsilno-sposterezhnomu kompleksi" [Image formation and processing in the sighting and surveillance complex]. Monohrafy. Lviv, NASV. 210 p. (with limited access). [in Ukrainian].

3. Khaustov D., Nastyshyn Yu., Khaustov Ya., Kyrychuk O., Stakh T., Malynych S. and Sidor R. (2023), "Suchasni prytsilno-sposterezhni kompleksy zrazkiv bronetankovoho ozbroiennia" [Modern sighting and surveillance systems for armored weapons]. Monohrafiia. Lviv, NASV. 310 p. ISBN 978-617-7689-08-8. [in Ukrainian].

4. Khaustov D., Kyrychuk O., Stakh T., Khaustov Ya., Sidor R., Burashnikov O., Ryzhov Ye.and Nastishin Yu. (2023), Target visibility index. Military Technical Collection. № 28. pp. 68-76. DOI: https://doi.org/10.33577/2312-4458.28.2023.68-76

5. Khaustov D., Kyrychuk O., Khaustov Ya., Stakh T., Zhyrna O. and Nastishin Yu. (2023), The measure of target saliency for target-oriented image fusion. *Scientific works of State Scientific Research Institute of AME TC*, Iss. 3(17). pp. 122-136, 202. DOI: 10.37701/dndivsovt.17.2023.15

6. T. M. Stakh, V. R. Bahan, O. A. Kyrychuk and etc. (2024), "Tema zakryta" [Topic closed]. *Military Technical Collection*. № 31 (t). NASV. Lviv. pp. 97-118. [in Ukrainian].

7. T. M. Stakh, R. I. Sidor, D. Ye. Khaustov, Ya. Ye. Khaustov, O. A. Kyrychuk, V. H. Mudryk and Yu. A. Nastyshyn. (2025), "Uzahalnena model zhyvuchosti zrazkiv bronetankovoho ozbroiennia i tekhniky" [Generalized survivability model of armored weapons and equipment samples]. *Military Technical Collection*. № 32. NASV. Lviv. pp. 28-43. [in Ukrainian].

8. Horbunov V. A. "Yeffektyvnost obnaruzhenyia tselei" [Target Detection Efficiency]. Vydavnytstvo Ministerstva oborony SRSR, M., 160 p. [in Russian].

9. Afanasev A. A. and Horbunov V. A. (1964), "Yeffektyvnost obnaruzhenyia tselei radyotekhnycheskymy smredstvamy nabliudenyia"[Target Detection Efficiency by Radio-Engineering Surveillance Equipment]. M. Voenyzdat. 122 p. [in Russian].

10. Huseinov A.B. and Perkov Y.E. (2005), Pokazately zametnosty letatelnykh apparatov y sposoby ykh snyzhenyia" [Aircraft Visibility Indicators and Methods for Reducing Them]. Moskva. Yzd-vo MAY. 96 p. [in Russian].

11. Johnson, J. (1958), Analysis of image forming systems. Technical report. U.S. Army Engineer Research and Development Laboratories. Available at: https: //home. cis.rit.edu/ ~cnspci/ references/johnson1958.pdf (Accessed July 11 2025).

12. Sjaardema, Tracy A.; Smith, Collin S.; Birch and Gabriel C. (2015), History and Evolution of the Johnson Criteria. SANDIA REPORT, SAND2015-6368. July 1 2015. https:// www. osti.gov/servlets/purl/1222446 (Accessed July 11 2025).

13. Vollmerhausen R.H. and Eddie Jacobs. (2004), The Targeting Task Performance (TTP) Metric. A New Model for Predicting Target Acquisition Performance. Technical Report AMSEL-NV-TR-230. Available at: https://apps.dtic. mil/dtic/tr/ fulltext/u2/a422493.pdf (Accessed July 11 2025).

14. D.Ye. Khaustov, O.O. Burashnikov, Ya.Ye. Khaustov and Yu. A. Nastyshyn. (2021), "Uzahalnena matematychna model vykonannia vohnevykh zadach ekipazhem tanka" [Generalized mathematical model of fire tasks performed by a tank crew]. *Ozbroiennia ta viiskova tekhnika*. №1(29). pp. 20-26. [in Ukrainian].

15. D.Ye. Khaustov, Yu.A. Nastyshyn and Ya.Ye. Khaustov. (2021), "Ymovirnist vykonannia vizualnoi zadachi yak syhmoidna funktsiia" [Probability of completing a visual task as a sigmoid function]. TsNDI OVT: *Ozbroiennia ta viiskova tekhnika*, No. 3 (31). pp. 80-94. [in Ukrainian].

16. Jonathan G. Hixson, Eddie L. Jacobs and Richard H. Vollmerhausen (2004), "Target detection cycle criteria when using the targeting task performance metric", Proc. SPIE 5612, Electro-Optical and Infrared Systems: Technology and Applications, (6 December 2004); https://doi.org/ 10.1117/12.577830

17. Approach to Conceptual Sensor Modeling. (2018), el. resource: Mattias Sonesson. A Probabilistic Dissertation, Linköpings Universitet, Linköping, Sweden, URL: http://liu.diva-portal.org/smash/record.jsf?pid=diva2%3A20633&dswid=-9203 (Accessed June 15 2025).

18. O.A. Kyrychuk, D.Ye. Khaustov, T.M. Stakh, Ya.Ye. Khaustov, R.I. Sidor, O. V. Zhyrna and Yu. A. Nastyshyn. (2024), "Tema zakryta" [Topic closed]. *Naukovi pratsi DNDI ViS OVT*. Vyp. 1(2)-T. Cherkasy. pp. 83-110. [in Ukrainian].

19. Contrast Threshold. International Dictionary for Marine Aids to Navigation. https://www.iala-aism.org/wiki/ dictionary/ index.php/Contrast_Threshold (Accessed July 16 2025).

20. Fisher R.A. (1925), Statistical Methods for Research Workers. Oliver and Boyd (Edinburgh), ISBN 978-0-05-002170-5.

21. O. A. Kyrychuk, D. Ye. Khaustov, T. M. Stakh, Ya. Ye. Khaustov, R. I. Sidor, A. V. Shulhin and Yu. A. Nastyshyn. (2023), "Tema zakryta" [Topic closed]. *Military Technical Collection*. № 29t. pp. 21-32. [in Ukrainian].

22. David Stump. (2014), Digital Cinematography. Fundamentals, Tools, Techniques, and Workflows (1st Edition), Routledge. New York. 498 p. DOI: https://doi.org/ 10.4324/ 9780240817927

23. Gregory Hollows and Nicholas James. Imaging Resource Guide. Edmund Optics: Resolution. Internet dzherelo: https:// www.edmundoptics.com/knowledge-center/application-notes/imaging/resolution/ (Accessed July 18 2025).

24. Hniedenko B. V. (2010), "Kurs teorii ymovirnostei" [Probability theory course]. Kyiv VPTs Kyivskyi universytet. 464 p. [in Ukrainian].

25. "Samokhidna haubytsia M109: instrument bohiv viiny" [M109 Self-Propelled Howitzer: Tool of the Gods of War]. *Militarnyi*, Internet vydannia https://militarnyi. com/uk/articles/

samohidna-gaubytsya-m109-instrument-bogiv-vijny/ (Accessed July 1* 2025). [in Ukrainian].

26. "T-72 : Stattia z Vikipedii " [T-72 : Wikipedia article]. https://uk.wikipedia. org/wiki/T-72 (data zvernennia 18.07.2025). [in Ukrainian].

27. T.M. Stakh, R.I. Sidor, V.P. Hladun, V.V. Sakharchuk, D.Ye. Khaustov, Ya.Ye. Khaustov, O.A. Kyrychuk, V.V. Lytvyn, O.R. Khobor, O.V. Zhyrna and Yu.A. Nastyshyn. (2025), "Kompiuterni zastosunky dlia avtomatyzatsii protsedury zboru danykh pro tsil z kamer BPLA v interesakh zrazkiv bronetankovoho ozbroiennia" [Computer applications for automating the procedure for collecting target data from UAV cameras in the interests of armored weapons samples*]. Zbirnyk naukovykh prats DNDI VS OVT,* Vyp. 2(24). pp. 126-145. DOI: https://doi.org/10.33577/2312-4458.32.2025.28-43. [in Ukrainian].

МЕТОД ОЦІНКИ ЙМОВІРНОСТЕЙ ВИЯВЛЕННЯ ТА РОЗПІЗНАВАННЯ ЦІЛЕЙ ЗА ЇХНІМИ ПРИКМЕТНИМИ РОЗМІРНИМИ ОЗНАКАМИ НА ЦИФРОВИХ ЗОБРАЖЕННЯХ З ПРИЦІЛЬНО-СПОСТЕРЕЖНОГО КОМПЛЕКСУ ЗРАЗКІВ БРОНЕТАНКОВОГО ОЗБРОЄННЯ

Т.М. Стах, Д.Є. Хаустов, В.Р. Баган, Я.Є. Хаустов, О.А. Киричук, Ю.А. Настишин

Стрімкий розвиток технологій комп'ютерного бачення та штучного інтелекту (ШІ) відкрив нові можливості для моніторингу поля бою у режимі реального часу, водночас створивши проблему ефективного відбору значущих електронних зображень серед гігантських масивів даних. У роботі розглянуто задачу автоматизованої оцінки якості відображення цілі на електронних зображеннях (ЯВЦнЕЗ) для підвищення ефективності виявлення та розпізнавання озброєння і військової техніки (ОВТ). На основі запропонованої системи індексів прикметності цілі за її розмірними ознаками та теоретично введених аналітичних залежностей ймовірностей виявлення та розпізнавання цілі від розмірів її геометричних сигнатур запропоновано метод виявлення та розпізнавання цілей за їхніми прикметними розмірними ознаками. Запропонований метод дозволяє кількісно вимірювати ЯВЦнЕЗ за її розмірними прикметними ознаками та ввести критерії успішності виконання візуальних задач виявлення та розпізнавання цілі на основі 50% порогового значення ймовірностей виявлення та розпізнавання. Показано, що лише ті зображення, які відповідають критеріям успішності виконання візуальних задач, можуть вважатися придатними для подальшої обробки засобами ШІ для розпізнавання цілей, що дозволяє знизити кількість хибних розпізнавань.

Особливу увагу приділено процесу розпізнавання як проміжному етапу збору інформації про ціль. Запропоновано використовувати геометричні прикметні ознаки зразків ОВТ як сигнатури їхнього розпізнавання. Теоретично введено аналітичну залежність ймовірності розпізнавання за розмірами сигнатур та визначено критерій успішності виконання візуальної задачі розпізнавання цілі. Такий підхід значно знижує кількість хибних ідентифікацій об'єктів на зображенні, що не забезпечують для ШІ достатню ЯВЦнЕЗ.

Практична реалізація методу продемонстрована на експериментальних даних, отриманих для зображень основного бойового танка Т-64БВ, зібраних у польових умовах. Зображення реєструвалися цифровою фотокамерою на різних відстанях до цілі. Для них визначено ймовірності виявлення за загальними розмірами цілі як об'єкта інтересу та розпізнавання за анізометрією форми і розмірами характерних сигнатур. Результати підтвердили узгодженість теоретично розроблених критеріїв із візуальним аналізом.

Практичне значення роботи полягає у створенні основ для автоматизованого відбору якісних електронних зображень без участі оператора, що дає змогу суттєво знизити обсяг даних, які обробляються, підвищити точність і швидкість роботи систем комп'ютерного бачення та зменшити кількість хибних розпізнавань. Перспективними напрямами подальших досліджень є розширення методики на тепловізійні зображення та використання цифрових зображень для визначення відстані до цілі за розмірами її сигнатур на цифрових зображеннях.

Ключові слова: *збір даних про ціль, цифрове зображення, обробка зображень, комплексування зображень, штучний інтелект, видимість, прикметність за певною ознакою, помітність цілі, виявлення, розпізнавання, ідентифікація, розвідка, об'єкт, безпілотний літальний апарат.*