UDC 004.942:517.9:519.6
DOI: 10.15587/2706-5448.2026.352493

Roman Havdulskyi,
Liubov Zhuravchak

INCLUSION AND HOLES DETECTION IN 3D OBJECTS USING NUMERICAL MODELING AND NEURAL NETWORKS

The object of research is a stationary heat-conduction process in three-dimensional heterogeneous media which components are in ideal thermal contact.

The addressed problem is the difficulties in detecting, localizing, and classifying inclusions in 3D objects. Most methods rely on employing active thermography which produces data as time-dependent sequences. Both capturing and processing it is quite costly and resource required.

The paper proposes a hybrid approach that relies on steady-heat thermograms, which are significantly easier and cheaper to capture. The developed approach uses one of the variants of the indirect near-boundary element method (INBEM), SOTA detectors (YOLO 11n and 12n), and a custom depth head based on ConvNeXt V2. Additionally, the paper showed that attention-centric architecture is promising for processing steady heat images.

INBEM with near-boundary elements in the form of families of points achieves an execution time of approximately 50 seconds per sample, with a general maximum error of approximately 0.08. This enabled the creation of a large dataset, comprising approximately 130K samples. Additionally, a testing dataset with a size of 7K and slightly different variance is obtained. Both YOLO 11n and 12n showed mAP50:95 metric results on the testing dataset of 85.2% and 89.4%, respectively. The precision/recall for both models are the following: 92.0/92.8 and 92.3/96.3. The depth head showed a MAPE of about 2%.

The proposed method focuses on detecting inclusions and holes using steady heat images, so it is suitable for relatively low-cost analysis, as obtaining such data is easier and quicker than collecting time-dependent data. It may be useful to screen slab-like structures, such as photovoltaic panels. Wall diagnostics is one possible future application area, as the work can be extended to semi-infinite objects. Thus, the results may serve as a basis for a low-cost inspection tool.

Keywords: *YOLO, steady-heat thermography, non-destructive testing, synthetic data generation, numerical modeling.*

Received: 13.11.2025
Received in revised form: 27.01.2026
Accepted: 18.02.2026
Published: 28.02.2026

How to cite
Havdulskyi, R., Zhuravchak, L. (2026). Inclusion and holes detection in 3D objects using numerical modeling and neural networks. Technology Audit and Production Reserves, 1 (2 (87)), 113–122. https://doi.org/10.15587/2706-5448.2026.352493

1. Introduction

Recognizing holes and inclusions in physical objects has significant value, as their presence in any medium alters the structure, signaling a degradation process or potential problem that has developed over time [1]. Thermography is an option for solving the problem of determining the internal structure [2]. It serves as a means to identify a defect [3, 4] or to verify the correctness of the internal structure [5]. As a resource-efficient solution in terms of cost, steady-heat thermography is considered a low-cost solution despite its recognition limitations and the time required [6] to reach steady heat distribution in the object. It's suitable for analyzing physical objects that are under the influence of a relatively steady heat source, such as the sun or any other artificial source. For instance, a solar panel [7, 8] or a building's wall placed under the open sun may achieve a steady heat state at a certain point. Thus, analyzing thermographic images of such objects may be considered a technique to verify the correctness of their internal structure [9, 10]. Furthermore, it's essential to distinguish between holes and inclusions, estimate their position (including depth) [11, 12], size, and physical properties, such as conductivity parameters, in the case of an inclusion.

Analyzing such thermographic images by neural network models, such as CNNs [13] or transformer-based models [14, 15], is an efficient way. However, such approaches require a lot of data to train on. Obtaining real experimental data is highly valuable. However, it is quite costly and time-consuming, particularly when capturing data in 3D [16]. Numerical methods, as tools for generating data, are widely used by the scientific community. Finite Element and Finite Difference Methods [17] are generally accurate, but they require discretizing the entire computational domain. While FEM handles complex geometries better thanks to flexible meshes, both methods can become challenging and computationally expensive when applied to intricate geometrical shapes. Boundary Element Method (BEM), compared to others, allows discretizing only the boundary and contact area, thus making meshing quite easier. However, BEM may be less accurate and slower compared to FEM, particularly for objects with multiple inclusions. An Indirect Near-Boundary Element Method (INBEM) stands as a solution to overcome some BEM's issues [18, 19], avoiding singularities and reducing implementation complexity [20, 21].

Thus, this paper proposes using the INBEM for 3D objects with inclusions and holes in an "express" implementation, which utilizes point families to replace integration over near-boundary elements

ISSN-L 2664-9969; E-ISSN 2706-5448

while maintaining moderate quality and accuracy [22, 23].

To locate and categorize the inclusions and holes, this paper proposes employing state-of-the-art models (YOLO 11 [24] and YOLO 12 [25]) from the computer vision sphere. These detectors enable extremely fast inference, making them suitable for rapid screening. Even if accuracy is not always perfect [26], they help build a balanced, scalable, and low-cost solution for large-scale inspections.

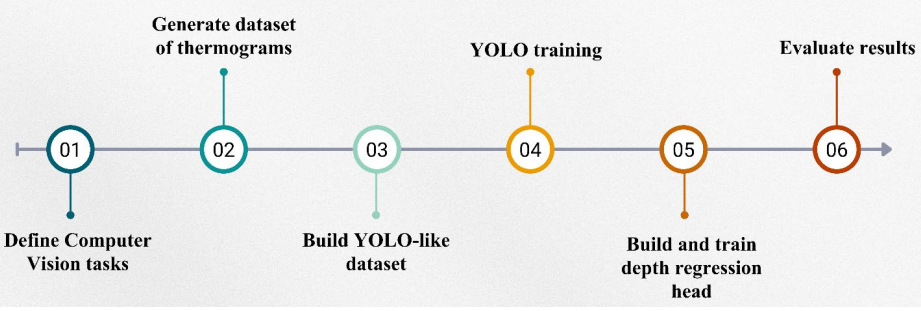


Fig. 1. Computer vision problem establishing process

Solving this task for detecting objects with a well-defined contour might be considered as quite relatively easy task. Detecting the inclusions or holes with different physical properties located at different depth levels just by considering a halo seems to be tricky and complex. Thus, the purpose of the paper is to verify the ability of such models to solve the task. Active thermography provides time-dependent [27] information, which makes more data available for analysis. Accordingly, a steady heat image provides less information. So, a regulated environment is employed. The variety of physical and geometrical properties is limited. The biases are enabled only for specific geometrical parameters, such as *x* and *y* position, keeping the same depth levels, sizes, and physical parameters. As these state-of-the-art (SOTA) models do not allow depth estimation, a separate custom depth head is considered. Despite some research aiming to estimate all geometrical parameters at once, this paper aims to employ SOTA ones as it allows switching them basically without any extra architecture change, so the detection head and depth head are treated independently.

Solving this task enables the more accurate detection of holes or inclusions and validation of the internal structure of 3D objects, such as solar panels, monolithic concrete blocks, or parallelepiped-shaped samples used in laboratory or industrial inspections.

The aim of research is to detect foreign inclusions and holes based on temperature or heat flux measured at the boundary of a three-dimensional object, caused by a stationary thermal field, to determine the geometric dimensions of inclusions (holes) and thermophysical characteristics using the indirect near-boundary method, where YOLO-based object detection is used to localize and classify defects on steady-heat thermograms, and a depth regression head is employed to estimate defect depth to support practical, rapid screening.

To achieve this aim, there are the following research objectives to complete:

1. To develop and evaluate a quick and simple method to generate datasets of steady heat thermograms by employing the INBEM.
2. To develop and evaluate a method for inclusion/hole detection and classification by employing state-of-the-art computer vision models, specifically YOLO 11 and YOLO 12.
3. To develop and evaluate a method for inclusion/hole depth estimation.

2. Materials and Methods

2.1. Object and subject of the research

The object of research is a stationary heat-conduction process in three-dimensional heterogeneous media which components are in ideal thermal contact.

The subject of research is neural network methods and tools for detecting and recognizing the characteristics of inclusions.

2.2. Problem and experiment establishing

To address the inverse problem, it is proposed to formulate it as a set of computer vision tasks. The overall process is presented in Fig. 1.

As shown in Fig. 1, there are 6 general steps:

1. The inverse problem is cast as object detection and its class recognition. Additionally, depth estimation is introduced as a regression task.
2. As these tasks require supervised learning, a dataset for training is required. Thus, the INBEM enables building a large dataset in a reasonable time, keeping both configurable object parameters and simulation accuracy.
3. As computer vision models, in general, work with images as input data, the thermographic signals are converted into images. The YOLO annotation is followed as well.
4. Both YOLO 11 and YOLO 12 must be trained to solve object detection and classification tasks, utilizing the obtained dataset.
5. To support YOLO results, a custom depth regression head is employed to process the halo around the detection as input and estimate its depth.
6. To evaluate performance, a separate dataset is required. Standard metrics as mAP, precision, recall, and F1 are used to assess object detection task performance. To assess the depth regression model, the following metrics are employed: MAE, MAPE, MSE, RMSE, and R^2.

Thus, the experiment is established in the following way:

1. The INBEM is enabled to generate a large number of thermographic samples with a wide variety of spatial and physical parameters. The 3D object sizes are $8 \times 4 \times 2$. The boundary conditions are the following: the bottom face is fixed at 0 degrees, the left, right, front, and back faces are insulated, and a heat flux of 10 is applied to the top face. The maximum accuracy error (methodology described further) is specified as 0.02 units.
2. All verified thermographic samples are preprocessed and converted into grayscale images. Additionally, data augmentation is employed to obtain samples with different levels of noise filtering and additive noise added.
3. Both YOLO 11 and YOLO 12 are to be trained on the object detection task.
4. Separately, the depth regression head is to be trained on the same dataset. To speed up training and evaluation, it uses ground truth bounding boxes (rather than YOLO-predicted detections) to extract the halo around the inclusion or hole. Also, the ground truth spatial information is provided as a separate input to the depth head.
5. Upon finishing training the models, they are to be evaluated employing already specified metrics.

2.3. Thermograms generation using INBEM

Before developing an approach to solving the inverse problem, it is necessary to obtain a series of datasets corresponding to the solutions of the direct problem with varying geometric and physical characteristics of the inclusion. The mathematical models of steady-state processes, constructed for methods based on the use of natural and artificial potential fields (gravitational, magnetic, electric, thermal, filtration) to detect inhomogeneities within an object, consist of Laplace equations

$$P_0^{(m)}(u^{(m)}(x)) = \Delta u^{(m)}(x) = 0,\ x \in \Omega_m,\ \Omega_m \subset R^3,\ m = 0.1, \quad (1)$$

supplemented with prescribed boundary conditions

$$u(x)=f_{\Gamma}^{(1)}(x), x\in\partial\Omega^{(1)}, \frac{\partial u(x)}{\partial \mathbf{n}(x)}=f_{\Gamma}^{(2)}(x), x\in\partial\Omega^{(2)}, \quad (2)$$

and ideal contact conditions at the interface between media

$$u^{(0)}(x)=u^{(1)}(x), \lambda_0\frac{\partial u^{(m)}(x)}{\partial \mathbf{n}^{(m)}(x)}=-\lambda_1\frac{\partial u^{(s)}(x)}{\partial \mathbf{n}^{(s)}(x)}, x\in\partial\Omega_1, \quad (3)$$

where Ω_m – a homogeneous domain with a constant physical property λ_m (conductivity); $\boldsymbol{n}^{(0)}(x)=(\boldsymbol{n}_1^{(0)}(x),\boldsymbol{n}_2^{(0)}(x),\boldsymbol{n}_3^{(0)}(x))$ – the unit outward normal vector uniquely defined on the boundary $\partial\Omega_0$; $\Delta=\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2}$ – the Laplace operator; $x=(x_1,x_2,x_3)$, $\cup_{q=1}^{2}\partial\Omega^{(q)}=$ $=\partial\Omega$; $\partial\Omega^{(1)}\cap\partial\Omega^{(2)}=\varnothing$, $\partial\Omega=\Omega_0\cup\Omega_1\cup\partial\Omega_1$, $\partial\Omega_1$ – the boundary of the inclusion Ω_1; and x_1,x_2,x_3 is the Cartesian coordinate system.

In the case of a hole, there is a problem (1), (2), where the second of conditions (2) is given on the boundary of the hole.

To solve the boundary value problem (1)–(3) or (1), (2), the indirect near-boundary element method was used, which has demonstrated higher accuracy compared to the boundary element method, as well as lower requirements for the programmer's mathematical expertise.

The external near-boundary domains are discretized into near-boundary elements (NBEs) G_v^m $(v=1,\ldots,V_m)$ [28, 29] and families of points G_v^{mk} $(k=1,\ldots,K_{mv})$ are introduced on them. For each G_v^m unknown functions $\varphi_v^{(m)}(x)$ are introduced, which describe the collective behavior of fictitious sources placed at the points G_v^{mk}. The visualization of the discretization in the near-boundary zone of the main object (matrix) and in the near-boundary zone along the interface media (Fig. 2).

Visualization of discretization in the near-boundary zone of the main object (matrix) and in the near-boundary zone along the interface between media is presented in Fig. 2.

Since for the Laplace operator $\boldsymbol{P}_0^{(m)}(u^{(m)}(x))$ the known fundamental solution (FS) $\breve{U}^{(m)}(x,\xi)$ exists, the integral representations of the solutions of equations (1) and their normal derivatives have the following form [30]:

$$u^{(m)}(x)=\mathbf{F}^{(m)}(x,\breve{U}^{(m)}),$$
$$\partial u^{(m)}(x)/\partial \mathbf{n}^{(m)}(x)=\mathbf{F}^{(m)}(x,\breve{Q}^{(m)}), \quad (4)$$

where:

$$\mathbf{F}^{(m)}(x,\Phi^{(m)})=$$
$$=\sum_{v=1}^{Vm}\sum_{k=1}^{Kmv}\Phi^{(m)}(x,\xi^{v0k})\varphi_v^{(m)}(\xi^{v0k}),$$
$$\breve{U}^{(m)}(x,\xi)=\frac{1}{4\pi\lambda_m r},$$
$$\breve{Q}^{(m)}(x,\xi)=\sum_{l=1}^{3}\breve{Q}_l(x,\xi)\mathbf{n}_l^{(m)}(x),$$
$$\breve{Q}_l^{(m)}(x,\xi)=-\frac{y_l}{4\pi r^3},$$

$|x|$ – the modulus of vector x,

$$\xi=(\xi_1,\xi_2,\xi_3)\in R^3,$$
$$r^2=\sum_{i=1}^{3}y_i^3, y_i=x_i-\xi_i.$$

By directing x in equation (4) from the interior of the domain Ω toward its boundary $\partial\Omega$ to satisfy conditions (2) and (3), let's obtain the boundary integral equations (BIE):

$$\mathbf{F}^{(0)}(x,\breve{U}^{(0)})=f_{\Gamma}^{(1)}(x), \quad x\in\partial\Omega^{(1)},$$
$$\mathbf{F}^{(0)}(x,\breve{Q}^{(0)})=f_{\Gamma}^{(2)}(x), \quad x\in\partial\Omega^{(2)}; \quad (5)$$

$$\mathbf{F}^{(0)}(x,\breve{U}^{(0)})-\mathbf{F}^{(1)}(x,\breve{U}^{(1)})=0,$$
$$\lambda_0\mathbf{F}^{(0)}(x,\breve{Q}^{(0)})+\lambda_1\mathbf{F}^{(1)}(x,\breve{Q}^{(1)})=0. \quad (6)$$

The boundary integral equations (5), (6), after approximating the functions $\varphi_v^{(m)}(x)$ by unknown constants d_v^m, are written in the form of a system of linear algebraic equations (SLAE):

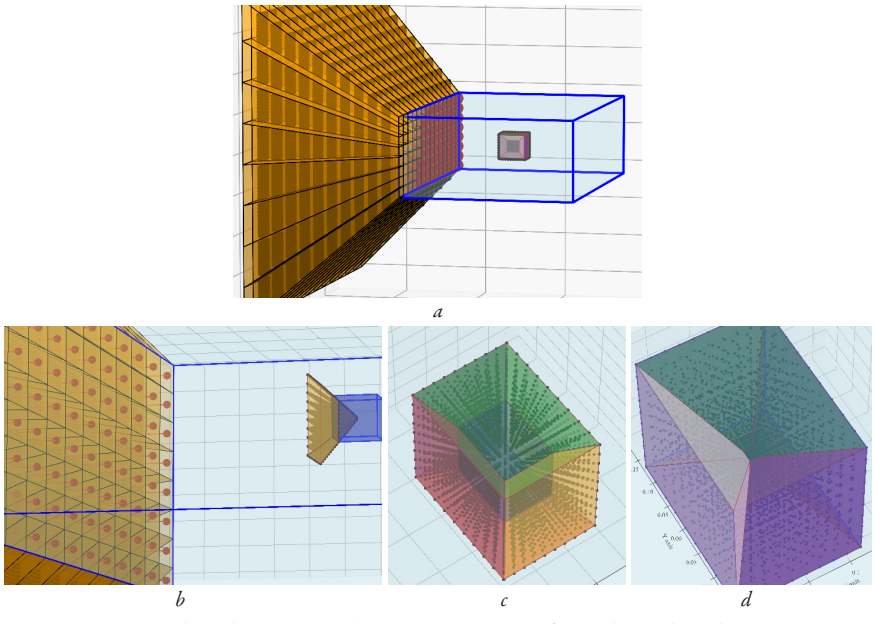
$$\sum_{v=1}^{V_0}\breve{A}_v^0(x^{0w},\breve{U}^{(0)})d_v^0=f_{\Gamma}^{(1)}(x^{0w}), x^{0w}\in\partial\Omega^{(1)}, w=1,\ldots,V_0^{(1)},$$
$$\sum_{v=1}^{V_0}\breve{A}_v^0(x^{w},\breve{Q}^{(0)})d_v^0=f_{\Gamma}^{(2)}(x^{0w}), x^{0w}\in\partial\Omega^{(2)}, w=V_0^{(1)}+1,\ldots,V_0^{(2)},$$
$$\sum_{v=1}^{V_0}\breve{A}_v^0(x^{1w},\breve{U}^{(0)})d_v^0-\sum_{v=1}^{V_1}\breve{A}_v^1(x^{1w},\breve{U}^{(1)})d_v^1=0, x^{1w}\in\partial\Omega_1, w=1,\ldots,V_1,$$
$$\lambda_0\sum_{v=1}^{V_0}\breve{A}_v^0(x^{1w},\breve{Q}^{(0)})d_v^0+\lambda_1\sum_{v=1}^{V_1}\breve{A}_v^1(x^{1w},\breve{Q}^{(1)})d_v^1=0, \quad (7)$$

where $\breve{A}_v^{(m)}(x,\Phi^{(m)})=\sum_{k=1}^{K_{mv}}\Phi^{(m)}(x,\xi^{v0k}), \breve{\Phi}\in\{U,Q\}$.

Having determined the unknowns d_v^m from (7), it is possible to use formulas analogous to (4) to compute the desired functions $u^{(m)}(x)$ and their normal derivatives at both internal points of each subdomain Ω_m, and at the interface between media, since the inclusions Ω_1 and the surrounding medium Ω_1 are now treated as completely independent domains:

$$u^{(m)}(x^{mz})=\sum_{v=1}^{V_m}\breve{A}_v^m(x^{mz},\breve{U}^{(m)})d_v^m+b^m(x^{mz},\breve{U}^{(m)}), x^{mz}\in\mathbf{Z}_m,$$
$$\frac{\partial u^{(m)}(x^{mz})}{\partial \mathbf{n}^{(m)}(x)}=\sum_{v=1}^{V_m}\breve{A}_v^m(x^{mz},\breve{Q}^{(m)})d_v^m+b^m(x^{mz},\breve{Q}^{(m)}), \quad (8)$$

where Z_m – the set of observation points $x^{mz}\in\partial\Omega_m\cup\Omega_m$.



a *b* *c* *d*

Fig. 2. Near-boundary zones visualization: *a* – overview of one side near-boundary zone; *b* – collocation points and interface zoomed; *c* – near-boundary zone of the inclusion as a part of the interface; *d* – near-boundary zone of the matrix directed into the hole/inclusion as a part of the interface

ISSN-L 2664-9969; E-ISSN 2706-5448

It is worth mentioning that the use of point families on NBEs made it possible, unlike in the BEM, to avoid prior analytical extraction of singularities (principal values) when computing the normal derivatives of the desired functions. Point families significantly simplify the problem-solving algorithm, as they replace integration over NBEs with the summation of products of the fundamental solution and the unknown (7) or known (8) source intensities. Thus, they can be recommended for obtaining initial approximations as a rapid method for solving inverse problems of mathematical physics, where time optimization is prioritized over accuracy.

2.4. Verification of simulation results

To verify the accuracy of the simulation, ten verification points are distributed along the diagonal line of each boundary face of the three-dimensional figure, and at these points, the error between the prescribed boundary conditions and the computed values (8) is evaluated. To ensure the correctness of the matrix – inclusion interface, the solution is additionally computed by (8) at every collocation point of the interface twice. Once by considering the point as belonging to the matrix and once as part of the inclusion, and the residual is defined as the absolute difference between the two values. The overall verification metric is then obtained by taking the maximum error over all boundary and interface verification points and comparing it with the tolerance parameter.

2.5. Converting thermographic signals into images

Once the simulation results are verified, the temperature distribution is calculated for the top side of the object. As it comes from (8), calculating the temperature distribution in high resolution requires a significant amount of time. Thus, low resolution mesh grid is created and calculated. To achieve high-resolution data, the mesh grid data is interpolated by the Clough-Tocher cubic method. The data in the form of a temperature distribution is baseline-corrected by subtracting the 5th percentile to remove background bias while preserving local contrast. Then the value above the 99th percentile is clipped and linearly normalized to the range of 0 to 1. Then, it is normalized into RGB format.

The formal representation is given:

$$\begin{aligned} T_{rel} &= T - percentile(T,5),\\ T_{clip} &= \min(T_{rel}, P_{99}),\\ T_{norm} &= \frac{T_{clip}}{P_{99}+10^{-6}},\\ T_{RGB} &= (255\cdot T_{norm}). \end{aligned} \tag{9}$$

The simulation results obtained using the "express" INBEM implementation might include computational noise not typical for such thermography images. It's brought by the discretization nature (noticeable when an inclusion/hole is placed far from the surface). Thus, the distribution is smoothed out by Gaussian filtering with its intensity specified by the sigma value. The resulting image is saved as a JPEG, and the dataset, including all metadata, is stored as a YOLO-style dataset.

2.6. Data augmentation

To extend it and ensure robustness against noise, a Gaussian noise augmentation strategy biased toward higher variance levels is implemented. Specifically, each image in the dataset was perturbed with additive Gaussian noise of standard deviation $\sigma \in [10, 50]$, where σ was sampled according to a Beta distribution parameterized to favor values near the upper bound of the interval. The weighted-high strategy, with a bias of 0.7, is employed to distribute the noise level evenly across images. The augmentation was applied consistently on an image-wise basis using Gaussian noise with fixed variance per sample. The resulting dataset contains both the original clean images and their noise-augmented counterparts, thereby providing a controlled yet noise-intensive extension of the baseline data. This dual structure is explicitly designed to evaluate and enhance the model's resilience to varying levels of sensor noise.

2.7. Dataset splitting

The training dataset is split as 80% for training and 20% for validation. The split logic is defined so that children of the same parent belong to the same split as the parent. The evaluation dataset is built similarly to the training dataset, but with a different variance of x and y coordinates, while maintaining the same variance of other parameters. The dataset batch is visualized in Fig. 3.

2.8. Building a depth regression head

The inclusion/hole bounding box is expected as the input. Then, a fixed-size area centered at the object coordinates is isolated. Normalized coordinates are converted to pixel values, and the patch boundaries are computed accordingly. If the region extends beyond image borders, symmetric zero-padding is applied to preserve the required patch dimensions without distorting thermal gradients. The patch dimension is specified as $W = H = 224$. The extracted image is normalized according to the ImageNet-specific [31] parameters listed in Table 1.

The structure of the depth head is provided in advance. The proposed neural network utilizes a lightweight convolutional backbone – ConvNeXt V2-pico/tiny [32] pretrained on the ImageNet dataset. Its learning rate and other parameters for transfer learning are chosen differently from those for the heads. Its classification head has been removed, and global average pooling is employed to obtain a vector of features with a dimension of $C = 768$. Thus, the backbone is used as a spatial and semantic information feature descriptor:

$$\begin{aligned} &f_{img} = Backbone(I) \in R^{BxC},\\ &I \in R^{Bx3xHxW}. \end{aligned} \tag{10}$$

However, considering only the patch features might not seem sufficient. Thus, contextual features, specifically the size and position of an inclusion/hole, are considered:

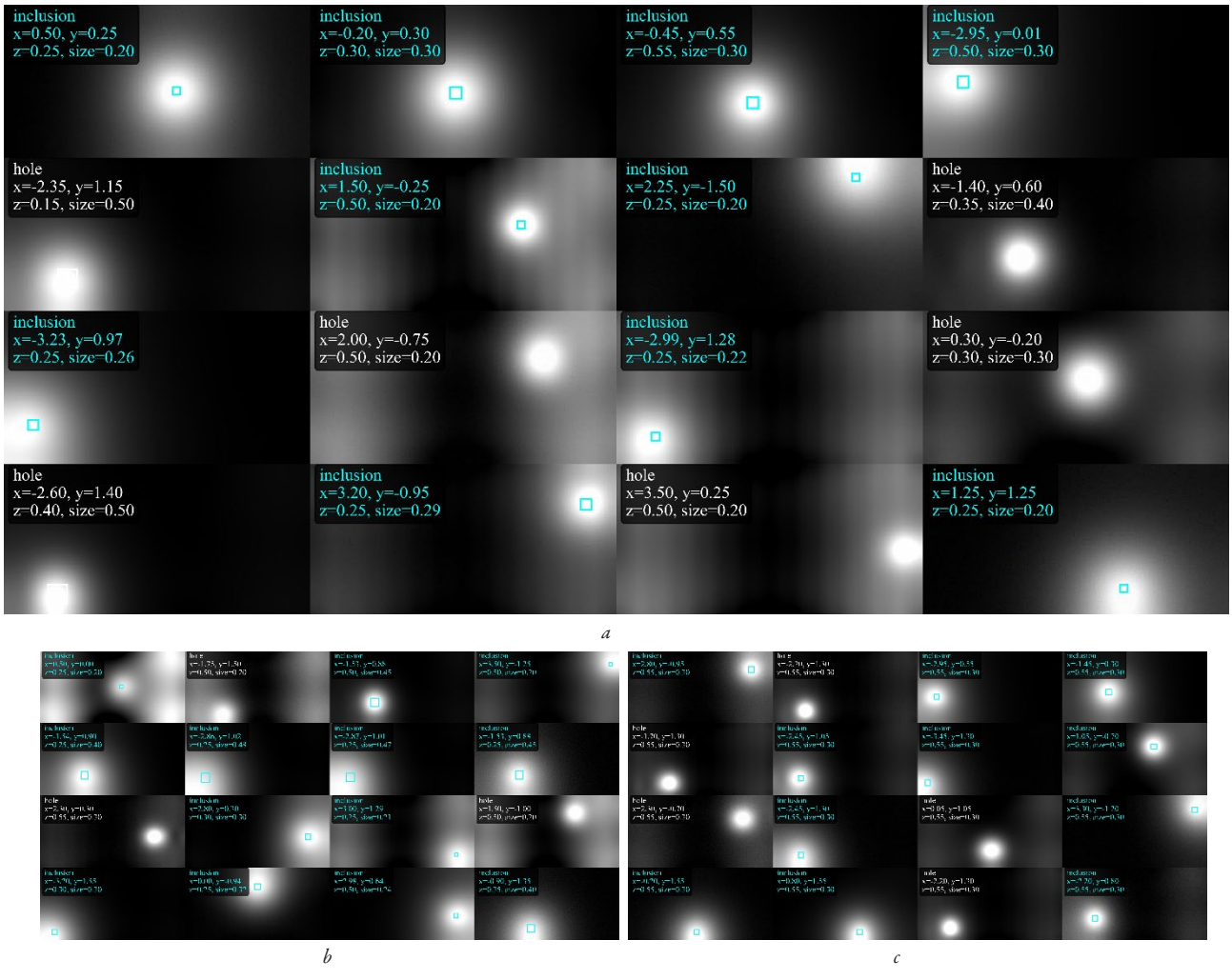
$$\begin{aligned} &f_{ctx} = Linear_2\left(SwiGLU\left(Linear_1\left(x_{ctx}\right)\right)\right) \in R^{Bxd_{out}},\\ &x_{ctx} \in R^{Bxd_{out}}, \end{aligned} \tag{11}$$

where $d_{ctx} = 2$ and $d_{out} = 32$.

As expressed in formula (9), SwiGLU [33] is used as the activation function. The layer, referred to as $Linear_1$, produces an output of dimension 64, which is processed by an activation function. $Linear_2$ then outputs contextual features of the specified dimension

$$y = Linear_2\left(Dropout\left(SwiGLU\left(Linear_1\left([f_{img}; f_{ctx}]\right)\right)\right)\right). \tag{12}$$

The concatenation of both semantic and contextual information is processed by the regression head, which consists of two linear layers with SwiGLU as the activation function and dropout regularization with a dropout probability of 0.05. The layer, referred to as $Linear_2$, produces an output of dimension size 256.



a

b *c*

Fig. 3. Dataset batches visualization: *a* – batch No. 1; *b* – batch No. 2; *c* – batch No. 3

Table 1

Normalization parameters

Parameter	*R*-value	*G*-value	*B*-value
mean	0.485	0.456	0.406
std	0.229	0.224	0.225

3. Results and Discussion

3.1. Generating datasets of steady heat thermograms using the INBEM

The INBEM is implemented using Python as a programming language, and NumPy is employed to build SLAE and solve it. To achieve optimal performance, vectorization [34] is employed.

A simulation is performed for a 3D object with the following dimensions: 8 × 4 × 2. A mesh of size 120 × 60 is chosen to extract temperature data using INBEM. Then it's interpolated up to shape 800 × 400. The following near-boundary zone split is chosen: 14 × 10 for the top side, 8 × 14 for the left and right sides, and 8 × 10 for the front and back sides of the matrix. Each near-boundary zone of the matrix includes 10 columns, 10 rows, and 10 depth levels. The min margin from the point to the boundary of each 3D zone is defined as 0.1 units. The boundary thickness is specified as 3. As for the inclusion/hole, each side has one near-boundary zone. Each near-boundary zone of the inclusion/hole includes 10 columns, 10 rows, and 10 depth levels. The min margin from the point to the boundary of each 3D zone is defined as 0.001 units. The thickness of the interface between media is specified by the formula

$$\text{interfaceThickness} = \min\left(\textit{inclusionSize}/2 - 0.01, 0.25\right). \quad (13)$$

Such a coordinated choice of simulation parameters is necessary to achieve a balance between execution time and the average level of accuracy. They are determined by empirically adjusting the thickness of near – boundary zones, the number of points, and their proximity to the boundary. It showed that an increasing number of them might result in achieving better results, but the execution time increases significantly (not linearly). Additionally, modifying those parameters may result in instability due to numerical overflow. The parameters provided are considered the most suitable, considering the average stability and accuracy achieved during empirical testing. Some attempts have been made to utilize well-known optimization techniques, but the parameters found appear to be generally inefficient, yielding quite good results only in specific use cases. Thus, considering the time required to run the optimization algorithm, the late approach does not appear to be particularly efficient and warrants further research.

The general size of the train split is 108 K, and of the validation split is 25K. The average maximum error rate is approximately 0.08, as measured by the method described in the previous chapter.

Both holes and inclusions are considered upon dataset generation. As for inclusions, conductivities such as 0.5, 2, and 4 are considered. The following sizes are considered: 0.2, 0.3, 0.4, 0.5. The variance

ISSN-L 2664-9969; E-ISSN 2706-5448

of other geometrical parameters for the training dataset is provided in Table 2, where x, y, z correspond to x_1, x_2, x_3.

Table 2

Dataset parameters

Parameter	min	max	step
Size 0.2			
x	–3.5	3.5	0.25
y	–1.5	1.5	0.25
z	–0.5	0.5	0.25
Size 0.3			
x	–3.45	3.55	0.25
y	–1.45	1.55	0.25
z	–0.45	0.55	0.25
Size 0.4			
x	–3.4	3.35	0.25
y	–1.4	1.35	0.25
z	–0.4	0.35	0.25
Size 0.5			
x	–3.35	3.4	0.25
y	–1.35	1.4	0.25
z	–0.35	0.4	0.25

To increase variance and mix data with those that are outside the distribution described, an additional set of elements is included. The inclusion/hole centers are arranged within narrow peripheral zones (0.8 units inward from each lateral boundary) along the four principal faces (left, right, top, and bottom), ensuring systematic coverage of regions adjacent to the object boundaries while maintaining minimal offsets of 0.5 and 0.4 units along the X, Y, and Z axes, respectively.

For each inclusion/hole size in the range of 0.20–0.50 (step 0.01), up to five uniformly spaced centers are generated along each edge band relatively to the top face. The centers are separated by a minimal gap proportional to the inclusion/hole size and perturbed by a small stochastic displacement to avoid strictly regular grid alignment. The distribution depths of the inclusion/hole are as follows: 0, ±0.25, and ±0.5.

All the items are processed as described in the previous chapter. The following sigma values are used as filtering parameters: 15 and 30. There is one exception: elements with a size of 0.2 and a depth of –0.5 are always processed using a 30 value of sigma. The purpose is to eliminate all the computational noise that is noticeable when an inclusion/hole influence is barely visible on the thermogram. The other reason is that the noise might be considered by the model as the real inclusion/hole. As the nature of this noise bears no resemblance to the one encountered while taking a thermography photo, such intense filtering (sigma 30) is employed for deeply located media.

To test the model's performance, a slightly different dataset is generated, which includes both holes and inclusions with a conductivity of 2. The other parameters are described in Table 3.

Table 3

Dataset parameters

Parameter	min	max	step
Size 0.2			
x	–3.4	3.25	0.35
y	–1.4	1.4	0.35
z	–0.5	0.5	0.35

3.2. Inclusion/hole detection and classification using YOLO 11n and YOLO 12n

Training of models is done on cloud hosting, specifically using RTX Ada 4000. The task chosen for training is detection. The number of epochs is 35, the batch size is 48, and the image size is 800. For both YOLO v11n and YOLO 12n, the other training parameters are stated in Table 4.

Table 4

Training parameters

Parameter	Value	Parameter	Value
Patience	5	Box	7.5
Optimizer	AdamW	Cls	0.5
Close mosaic	5	Dfl	1.5
LR 0	0.001	Hsv_h, hsv_s, hsv_v	0
LR F	0.01	Degrees	0
Momentum	0.9	Translate	0.05
Weight decay	0.01	Scale	0.15
Warmup epochs	3	Shear	0.0
Warmup momentum	0.8	Perspective	0
Warmup bias lr	0.1	Mosaic	1

The best candidates of both YOLO and depth models, validated on the validation dataset split, are saved for further testing.

To evaluate the single-shot detectors' results, the following metrics are considered: precision, recall, mAP50, and mAP50-95:

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{14}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{15}$$

$$mAP_{50} = \frac{1}{C}\sum_{C=1}^{C} AP_C^{(IoU \geq 0.50)}, \tag{16}$$

$$mAP_{[0.50:0.95]} = \frac{1}{10}\sum_{t=0.50}^{0.95} mAP@t, \tag{17}$$

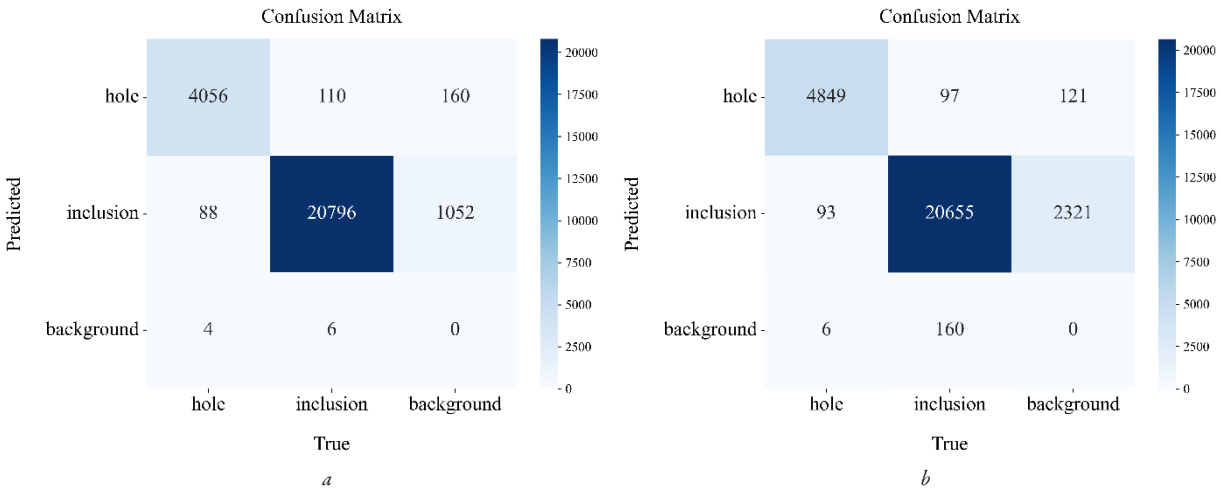
where IoU – intersection over union, AP_c – the average precision of each class, and C – the number of classes, which equals 2.

Both YOLO 11n and 12n showed the best performance on the validation split at the 30th epoch, just before mosaic was turned off. The result of validating the best candidate is presented in Table 5 and Fig. 4.

Table 5

YOLO validation results

Class	Instances	P	R	mAP50	mAP50-95
YOLO 11n					
All	25860	98.3	99.1	99.4	96.9
Hole	4948	97.5	99.1	99.3	97.3
Inclusion	20912	99.0	99.0	99.5	96.5
YOLO 12n					
All	25860	97.5	97.5	99.3	96.5
Hole	4948	98.3	98.5	99.4	97.8
Inclusion	20912	96.7	96.5	99.2	95.3



a *b*

Fig. 4. Confusion matrix on the training dataset of: *a* – YOLO v11n; *b* – YOLO v12n

Additionally, each model has been tested out on the testing dataset, and the results are presented in Table 6 and Fig. 5.

Table 6

YOLO testing results

Class	Instances	*P*	*R*	mAP50	mAP50-95
YOLO v11n					
All	3592	92.0	92.8	96.8	85.2
Hole	1800	97.2	91.4	98.7	82.2
Inclusion	1792	86.7	94.2	94.9	88.3
YOLO v12n					
All	3592	92.3	96.3	98.6	89.4
Hole	1800	95.7	94.8	99.0	86.5
Inclusion	1792	88.8	97.7	98.2	92.3

3.3. Inclusion/hole depth estimation using a custom depth head and ConvNeXt V2 as backbone

To train the depth head, the AdamW optimizer is chosen, and the OneCycleLR scheduler is chosen to mimic the YOLO training procedure. The number of epochs is 100, and the size of the batch is 64. The loss criterion is chosen as L1Loss. Additional training parameters are listed in Table 7.

To evaluate the depth model's performance, the following metrics are considered: *MSE*, *MAE*, *MAPE*, *RMSE*, and R^2:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \widehat{y_i}\right)^2, \tag{18}$$

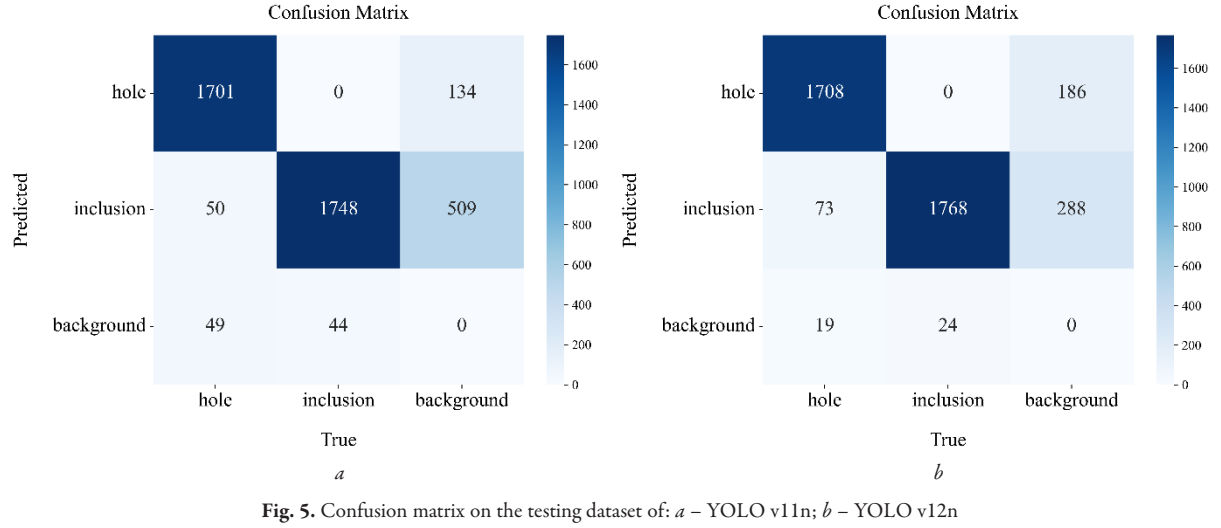
$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \widehat{y_i}\right|, \tag{19}$$

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{y_i - \widehat{y_i}}{yi}\right|, \tag{20}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \widehat{y_i}\right)^2}, \tag{21}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \widehat{y_i}\right)^2}{\sum_{i=1}^{N}\left(y_i - \overline{y_i}\right)^2}. \tag{22}$$

As the depth head is trained using two versions of ConvNeXt V2: Pico and Tiny, the validation metrics are presented in Table 8. The MSE metric is eliminated for presentation purposes, as it's smaller than 10^{-4}. The Pico version achieved best at the 100th epoch, and the Tiny – at the 97th epoch.



a *b*

Fig. 5. Confusion matrix on the testing dataset of: *a* – YOLO v11n; *b* – YOLO v12n

ISSN-L 2664-9969; E-ISSN 2706-5448

Table 7

Training parameters

Head		Backbone	
Parameter	Value	Parameter	Value
LR 0	0.001	LR 0	0.0001
Weight decay	0.0001	Weight decay	0.05
Final div factor	10000	Final div factor	10000
Div factor	25	Div factor	25
Anneal strategy	cos	Anneal strategy	cos
PCT start	0.05	PCT start	0.05

Table 8

Depth models' validation metrics

Backbone	*MAE*	*MAPE*, %	*RMSE*	R^2, %
ConvNeXt V2-pico	0.0006	0.3099	0.0069	99.96
ConvNeXt V2-tiny	0.0004	0.1966	0.0060	99.97

Additionally, each depth prediction model has been tested on the testing dataset, and the results are presented in Table 9.

Table 9

Depth models' testing metrics

Backbone	*MSE*	*MAE*	*MAPE*, %	*RMSE*	R^2, %
ConvNeXt V2-pico	0.0092	0.0285	2.4086	0.0958	93.44
ConvNeXt V2-tiny	0.0046	0.0212	2.3253	0.0676	96.73

It is worth mentioning that both YOLO models showed promising results for thermographic images, particularly when considering 3D objects with holes and inclusions. As stated in Table 6, both models perform quite well, showing mAP50:95 above 95% on the validation split of the training dataset. Surprisingly, YOLO 12n showed worse performance than 11n by about 1–2%. However, it might be evidence that 12n tends to slower learning and might be less likely to be overfit. Considering the results in Table 7, the following conclusions must be stated:

1. Localizing not just the halo but the inclusion/hole itself is proven to be promising.

2. YOLO 12n generalizes better by about 4%, considering just mAP50-95, while also showing significant improvements of recall and mAP50. At the same time, precision remains almost the same, showing only a slight improvement.

3. Such improvements shown by YOLO 12n bring an idea of further investigation of attention modules and other changes applied to its architecture, as might be promising in thermography analysis.

4. As expected, inclusions with different conductivities appear to be a tougher challenge for the model. However, the performance drop is not as significant as might be expected, considering the complexity of the dataset.

5. The depth prediction head showed quite outstanding performance on the testing dataset, which confirms the reasonability of its use in pair with the YOLO model to detect and classify the inclusion. Additionally, analyzing the patch that includes the halo and the surrounding area has been proven to be efficient.

As for the practical usage, the proposed method may be summarized as follows:

1. The combination of YOLO with INBEM enables fast yet reliable defect screening. Although both YOLO and INBEM are computationally efficient, the overall pipeline may not always be sufficiently accurate under certain conditions. The proposed approach, however, provides a controllable method that balances speed and accuracy.

2. In practical implementation, the researchers consider physical objects similar to PV panels, concrete walls, or any other with quite similar geometry. Thus, the results obtained in the paper may be used to continue researching such real physical objects in terms of identifying cracks, holes, and holes filled with materials such as water or similar substances. Then, employing transfer learning might be considered to achieve a promising result on a limited real dataset.

Let's compare the results achieved with the latest research on a similar topic [10]. It must be stated that the paper used for comparison employs non-steady heat images. Thus, the reference paper utilizes data that includes more valuable information for analysis. Considering the properties of the dataset elements, it is worth noting that they are quite similar, as the media are located at different depth levels; however, the media thickness and shape differ significantly. Additionally, each sample in this research includes multiple media, whereas our research uses a dataset with one inclusion/hole per sample. However, it is worth noting that our research encompasses defects of various types (inclusions with differing conductivity parameters and holes). The reference paper includes real data samples with an augmentation technique applied. To establish all the distinctions between our research and the reference one, let's proceed with the actual metrics comparison. According to the mAP (also referred to as AP in the reference paper but actually equivalent to mAP in our research context), the reference one yields the best result at 97.9% (YOLO-NAS-l), while our research achieves a metric of 98.6%. As our research and the comparison one use different classes for classification, the direct comparison might not make so much sense. However, it's still worth mentioning that the reference one focuses on an exact defect comparison, whereas our paper aims to determine whether the found one is a hole or an inclusion. To roughly compare the classification results, let's note that the reference one shows precision and recall at approximately 99% and 98%, respectively. According to our paper, the results are approximately 92% and 96%. Comparing those does not make so much sense due to their different natures and aims. But it is worth mentioning that the results achieved are quite high. The difference between the reference one and ours appears mainly due to employing steady heat thermograms in this paper, which is less informative. Also, the media differ in nature. Thus, the results obtained are considered quite promising, as they enable high performance while not requiring numerous additional tools and methods to collect data. As for the steady heat, the data might be captured relatively easily.

As for depth results, another paper [13] is considered. Basically, the difference between our research and reference one is almost the same as for the previous case. The major difference is that it employs cooling sequence processing to estimate parameters. The results of depth estimation achieved an MAPE of about 5.3% for most cases. Our research shows an MAPE of 2.3%, which is more than twice as good. Although the reference data obtained from real experiments and the dataset properties differ, our results prove that it's suitable for further research and is considered quite promising.

Considering the limitations found during the research, the following conclusions might be stated:

1. It's clear that analyzing steady heat images might result in recovering less information. For instance, different inclusions/holes with varying parameters within the matrix may result in quite similar thermographic images. Thus, extracting all possible parameters (size, depth, and conductivity) without seriously limiting the scope of parameters is considered a challenging and sometimes almost impossible task.

2. However, if the scope is limited and the possible variance is quite known with a bit of deviation, then steady heat analysis is considered a low-cost and efficient technique. Results obtained in the paper prove the point that state-of-the-art models, such as YOLO, are suitable for extracting spatial information for 3D figures with media inside.

3. INBEM in the "express" implementation form requires fine-tuning to be quite accurate and time-efficient method to simulate data.

Its fine-tuning requires some time and prior knowledge to achieve the required level of accuracy for some specific use cases. Additionally, the presence of high computational noise requires using smoothing techniques, which may still impact the final image, making it less natural. Alternatively, more accurate discretization techniques [35] might be applied. However, it leads to losing the simplicity of the method.

As for future research, such prospects might be considered:

1. As INBEM allows solving the same problem but for half-infinite objects, it might be more valuable for practical usage, as a building's walls might be considered semi-infinite as well. Thus, it appears to be a real use case for detecting such inclusions and holes in walls, which seem quite useful from a practical standpoint.
2. As YOLO 12n showed quite significant improvements, it seems to be reasonable proof that modifications that differ YOLO 12 from YOLO 11 are promising to employ in feature researchers considering both existing backbones with such modules and applying them to new ones.
3. As YOLO 26 [36] comes soon, it might be wise to check performance also against this model.
4. As our approach lacks the ability to extract conductivity parameters, it must be considered as a further step of the research.
5. Applying transformer-based solutions for both spatial reconstruction tasks and depth estimation might be reasonable, considering the current research state.

4. Conclusions

1. Using the "express" implementation of INBEM may be considered a reasonable and efficient method for obtaining a dataset of thermograms for physical 3D objects with an inclusion/hole inside. An execution time of approximately 50 seconds and an average maximum error of approximately 0.08 are achieved. Thus, both a relatively small error and a quite small execution time allowed the generation of a large dataset for this research. At the same time, it remains an open question about how those parameters will change in the case of using different, more complex discretization techniques that are more sophisticated in terms of expected accuracy. Additionally, to clearly state the advantages of INBEM, it should be compared with similar methods, which may be a point for further research. Therefore, the "express" INBEM is a promising method for further research as it allows solving direct problem relatively easily while keeping moderate and controllable performance.
2. The proposed method detects and classifies defects by employing state-of-the-art models as YOLO. Both YOLO 11n and 12n, employed in the research, have shown outstanding performance in terms of accuracy. Metrics such as mAP50, mAP50:95, precision, and recall demonstrated that the model generalizes well to thermography images and can locate inclusions and holes placed at different depth levels with high precision. In general, the best performer (YOLO 12n) showed metrics on the testing dataset: a precision of 92.3%, a recall of 96.3%, a mAP50 of 98.6%, and a mAP50:95 of 89.4%. It's worth mentioning that YOLO 12n showed a bit worse result on the validation dataset than YOLO 11n, but performed much better on the testing data than YOLO 11n. The improvement in performance is almost 5%. These findings reveal great potential of attention-centric architecture for thermogram analysis, as of impact on further scientific research. The proposed method may allow reaching quite accurate detection (compared to the ones for active thermography based) based just on the steady heat images. As for future research, verification on real data is necessary to prove it. Also, extending this research into semi-infinite objects allows for inspecting buildings' walls.
3. The proposed method to analyze the depth by employing the custom depth head has shown quite promising results. The depth head with ConvNeXt V2 as the backbone showed promising performance on both the validation and testing datasets, with an MAE that does not exceed 0.03 units. As of MAPE, the model achieved a result of approximately 2.3%. Further research should focus on integrating both the detector model and the depth head to measure their combined performance as a solid hybrid method. This result shows the potential for depth prediction in real 3D objects just by analyzing a steady heat image (considering the limitations, such as having defects of certain types, keeping boundary conditions, and so on).

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Financing

The research was performed without financial support.

Data availability

Data will be made available on reasonable request.

Use of artificial intelligence

The authors used artificial intelligence tools within acceptable limits. ChatGPT-5.2 and Grammarly were used only for language editing, improving translation quality, and refining the wording and structure of individual paragraphs.

The authors bear full responsibility for the final manuscript.

Declaration submitted by Roman Havdulskyi.

Authors' contributions

Roman Havdulskyi: Conceptualization, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft; ***Liubov Zhuravchak***: Methodology, Resources, Writing – review and editing, Supervision, Project administration.

References

1. Rezayiye, R. K., Ibarra-Castanedo, C., Maldague, X. (2024). Methods for Corrosion Detection in Pipes Using Thermography: A Case Study on Synthetic Datasets. *Algorithms, 17 (10),* 439. https://doi.org/10.3390/a17100439
2. Salah, M., Werghi, N., Svetinovic, D., Abdulrahman, Y. (2025). *Multi-Modal Attention Networks for Enhanced Segmentation and Depth Estimation of Subsurface Defects in Pulse Thermography.* ArXiv. https://doi.org/10.48550/arxiv.2501.09994
3. Sels, S., Bogaerts, B., Verspeek, S., Ribbens, B., Steenackers, G., Penne, R. et al. (2020). 3D Defect detection using weighted principal component thermography. *Optics and Lasers in Engineering, 128,* 106039. https://doi.org/10.1016/j.optlaseng.2020.106039
4. Dudzik, S. (2015). Two-stage neural algorithm for defect detection and characterization uses an active thermography. *Infrared Physics & Technology, 71,* 187–197. https://doi.org/10.1016/j.infrared.2015.03.003
5. Szymanik, B., Psuj, G., Hashemi, M., Lopato, P. (2021). Detection and Identification of Defects in 3D-Printed Dielectric Structures via Thermographic Inspection and Deep Neural Networks. *Materials, 14 (15),* 4168. https://doi.org/10.3390/ma14154168
6. Kumar, R., Puranik, V. E., Gupta, R. (2024). Unveiling the Potential of Infrared Thermography in Quantitative Investigation of Potential-Induced Degradation in Crystalline Silicon PV Module. *Solar Energy Advances, 4,* 100049. https://doi.org/10.1016/j.seja.2023.100049
7. Spajić, M., Talajić, M., Mršić, L. (2024). Using CNNs for Photovoltaic Panel Defect Detection via Infrared Thermography to Support Industry 4.0. *Business Systems Research Journal, 15 (1),* 45–66. https://doi.org/10.2478/bsrj-2024-0003
8. Di Tommaso, A., Betti, A., Fontanelli, G., Michelozzi, B. (2022). A multi-stage model based on YOLOv3 for defect detection in PV panels based on IR and visible imaging by unmanned aerial vehicle. *Renewable Energy, 193,* 941–962. https://doi.org/10.1016/j.renene.2022.04.046

ISSN-L 2664-9969; E-ISSN 2706-5448

9. Ta, Q. T., Mac, V. H., Huh, J., Yim, H. J., Lee, G. (2025). Automatic detection of subsurface defects in concrete structures using state-of-the-art deep learning-based object detectors on the infrared dataset. *Engineering Structures, 329,* 119829. https://doi.org/10.1016/j.engstruct.2025.119829

10. Jeon, M., Yoo, S., Kim, S.-W. (2022). A Contactless PCBA Defect Detection Method: Convolutional Neural Networks With Thermographic Images. *IEEE Transactions on Components, Packaging and Manufacturing Technology, 12 (3),* 489–501. https://doi.org/10.1109/tcpmt.2022.3147319

11. Liu, J., Long, X., Jiang, C., Liao, W. (2024). Multi-feature vision transformer for automatic defect detection and quantification in composites using thermography. *NDT & E International,* 143, 103033. https://doi.org/10.1016/j.ndteint.2023.103033

12. Wu, Z., Chen, S., Feng, F., Qi, J., Feng, L., Tao, N. et al. (2023). Automatic defect detection and three-dimensional reconstruction from pulsed thermography images based on a bidirectional long-short term memory network. *Engineering Applications of Artificial Intelligence, 124,* 106574. https://doi.org/10.1016/j.engappai.2023.106574

13. Resendiz-Ochoa, E., Trejo-Chavez, O., Saucedo-Dorantes, J. J., Morales-Hernandez, L. A., Cruz-Albarran, I. A. (2024). Application of Thermography and Convolutional Neural Network to Diagnose Mechanical Faults in Induction Motors and Gearbox Wear. *Applied System Innovation, 7 (6),* 123. https://doi.org/10.3390/asi7060123

14. Era, I. Z., Zhou, F., Raihan, A. S., Ahmed, I., Abul-Haj, A., Craig, J. et al. (2024). *In-Situ Melt Pool Characterization via Thermal Imaging for Defect Detection in Directed Energy Deposition Using Vision Transformers*. ArXiv. https://doi.org/10.48550/arxiv.2411.12028

15. Zheng, S., Zhang, J., Zu, R., Li, Y. (2024). Vision transformer-enhanced thermal anomaly detection in building facades through fusion of thermal and visible imagery. *Journal of Asian Architecture and Building Engineering, 24 (4),* 2854–2868. https://doi.org/10.1080/13467581.2024.2379866

16. Campione, I., Lucchi, F., Santopuoli, N., Seccia, L. (2020). 3D Thermal Imaging System with Decoupled Acquisition for Industrial and Cultural Heritage Applications. *Applied Sciences, 10 (3),* 828. https://doi.org/10.3390/app10030828

17. Sabathiel, S., Sanchis-Alepuz, H., Wilson, A. S., Reynvaan, J., Stipsitz, M. (2024). Neural network-based reconstruction of steady-state temperature systems with unknown material composition. *Scientific Reports, 14 (1).* https://doi.org/10.1038/s41598-024-73380-1

18. Zhuravchak, L. M., Zabrods'ka, N. V. (2010). Nonstationary thermal fields in inhomogeneous materials with nonlinear behavior of the components. *Materials Science, 46 (1),* 36–46. https://doi.org/10.1007/s11003-010-9261-9

19. Zhuravchak, L., Kruk, O. (2015). Consideration of the nonlinear behavior of environmental material and a three-dimensional internal heat sources in mathematical modeling of heat conduction. *Mathematical Modeling and Computing, 2 (1),* 107–113. https://doi.org/10.23939/mmc2015.01.107

20. Zhuravchak, L. (2019). Computation of pressure change in piecewise-homogeneous reservoir for elastic regime by indirect near-boundary element method. *2019 IEEE 14th International Conference on Computer Sciences and Information Technologies (CSIT),* 141–144. https://doi.org/10.1109/stc-csit.2019.8929820

21. Zhuravchak, L.; Shakhovska, N., Medykovskyy, M. O. (Eds.) (2019). Mathematical Modelling of Non-stationary Processes in the Piecewise-Homogeneous Domains by Near-Boundary Element Method. *Advances in Intelligent Systems and Computing IV*. Cham: Springer, 64–77. https://doi.org/10.1007/978-3-030-33695-0_6

22. Havdulskyi, R., Zhuravchak, L. (2025). Recognition of inclusion characteristics using neural network methods in stationary process modeling. *Visnyk Natsionalnoho universytetu "Lvivska politekhnika". Seriia Informatsiini systemy ta merezhi, 17,* 75–92. https://doi.org/10.23939/sisn2025.17.075

23. Havdulskyi, R., Zhuravchak, L., Yakovyna, V.; Hovorushchenko, T., Savenko, O., Popov, P. T., Lysenko, S. (Eds.) (2025). Inclusion parameters estimation using INBEM and CNNs with gradient-based feature extraction. *2nd International Workshop on Intelligent & CyberPhysical Systems (ICyberPhyS 2025). CEUR Workshop Proceedings,* 4013. Available at: https://ceur-ws.org/Vol-4013/paper2.pdf

24. Khanam, R., Hussain, M. (2024). *YOLOv11: An Overview of the Key Architectural Enhancements*. ArXiv. https://doi.org/10.48550/arxiv.2410.17725

25. Tian, Y., Ye, Q., Doermann, D. (2025). *YOLOv12: Attention-Centric Real-Time Object Detectors*. ArXiv. https://doi.org/10.48550/arxiv.2502.12524

26. Sharma, A., Kumar, V., Longchamps, L. (2024). Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and Faster R-CNN models for detection of multiple weed species. *Smart Agricultural Technology, 9,* 100648. https://doi.org/10.1016/j.atech.2024.100648

27. Jaeger, B. E., Schmid, S., Grosse, C. U., Gögelein, A., Elischberger, F. (2022). Infrared Thermal Imaging-Based Turbine Blade Crack Classification Using Deep Learning. *Journal of Nondestructive Evaluation, 41 (4).* https://doi.org/10.1007/s10921-022-00907-9

28. Zhuravchak, L. M., Zabrodska, N. V. (2020). Using of partly-boundary elements as a version of the indirect near-boundary element method for potential field modeling. *Mathematical Modeling and Computing, 8 (1),* 1–10. https://doi.org/10.23939/mmc2021.01.001

29. Zhuravchak, L. M. (2024). Potential field modeling by combination of near-boundary and contact elements with non-classical finite differences in a heterogeneous medium. *Mathematical Modeling and Computing, 11 (2),* 373–384. https://doi.org/10.23939/mmc2024.02.373

30. Zhuravchak, L., Zabrodska, N. (2024). Algorithm for determining inclusion parameters in solving inverse problems of geoelectrical exploration using the profiling method. *Geodynamics, 1 (36),* 98–107. https://doi.org/10.23939/jgd2024.01.098

31. Russakovsky, O., Deng, J., Su, H., Krause, J., Sanjeev Satheesh, Ma, S. et al. (2014). *ImageNet Large Scale Visual Recognition Challenge*. ArXiv. https://doi.org/10.48550/arxiv.1409.0575

32. Woo, S., Debnath, S., Hu, R., Chen, X., Liu, Z., Kweon, I. S. et al. (2023). *ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders*. ArXiv. https://doi.org/10.48550/arxiv.2301.00808

33. Shazeer, N. (2020). *GLU Variants Improve Transformer*. ArXiv. https://doi.org/10.48550/arXiv.2002.05202

34. Arora, I. (2024). Improving Performance of Data Science Applications in Python. *Indian Journal Of Science And Technology, 17 (24),* 2499–2507. https://doi.org/10.17485/ijst/v17i24.914

35. Zhuravchak, L. (2023). Computational Aspects of the Use of Different Types of Near-boundary elements in Modeling the Non-stationary Heat Conduction Process. *2023 IEEE 18th International Conference on Computer Science and Information Technologies (CSIT).* Lviv: IEEE, 1–5. https://doi.org/10.1109/csit61576.2023.10324216

36. Sapkota, R., Harsha, C. R., Sharda, A., Karkee, M. (2025). *YOLO26: Key Architectural Enhancements and Performance Benchmarking for Real-Time Object Detection*. ArXiv. https://doi.org/10.48550/arxiv.2509.25164

✉ ***Roman Havdulskyi,*** *PhD Student, Department of Software, Lviv Polytechnic National University, Lviv, Ukraine, e-mail: roman.i.havdulskyi@lpnu.ua, ORCID: https://orcid.org/0009-0004-6737-6018*

Liubov Zhuravchak, *Doctor of Technical Sciences, Senior Researcher, Professor, Department of Software, Lviv Polytechnic National University, Lviv, Ukraine, ORCID: https://orcid.org/0000-0002-1444-5882*

✉ *Corresponding author*