SECTION 9.
INFORMATION TECHNOLOGIES AND SYSTEMS

SCIENTIFIC RESEARCH GROUP:

Didur Volodymyr Oleksandrovych
Bachelor student of Computer Science Department
Khmelnytskyi National University, Ukraine

Molchanova Maryna Oleksyivna
Teacher of Computer Science Department
Khmelnytskyi National University, Ukraine

Tyshchenko Olena Oleksandrivna
Teacher of Computer Science Department
Khmelnytskyi National University, Ukraine

Mazurets Oleksandr Viktorovych
Ph.D. in Technical Sciences, Associate Professor,
Associate Professor of Computer Science Department
Khmelnytskyi National University, Ukraine

APPROACH FOR COMPARATIVE ANALYSIS OF EFFECTIVENESS OF USING MOBILENETV3 AND VIT NEURAL NETWORK MODELS FOR GRAPHICAL LOCALIZATION OF DESTROYED BUILDINGS REMAINS AREAS

In the wake of armed conflicts, natural disasters, and large-scale industrial accidents, the rapid and accurate assessment of structural damage has become an urgent priority for humanitarian agencies, urban planners, and emergency responders alike. By contrast, automated image-based approaches promise to deliver timely, objective, and scalable insights into the extent and severity of destruction. Within this landscape, the graphical localization of destroyed building remains – i.e., the precise mapping of debris outlines, collapsed walls, and compromised structural elements – stands out as a critical capability for prioritizing rescue operations, planning reconstruction, and monitoring ongoing threats to public safety [1].

Over recent years, the emergence of lightweight convolutional architectures such as MobileNetV3 and the advent of transformer-based vision models (ViT) have opened new frontiers in on-device inference and high-accuracy scene understanding. MobileNetV3's efficient depth wise separable convolutions and neural architecture search optimizations make it particularly well suited for deployment on drones, ground robots, or other edge devices with limited computational budgets.

Despite their individual strengths, a systematic comparison of these two paradigms – mobile-optimized CNNs versus transformer models – remains underexplored in the specific context of disaster-zone imagery [2, 3]. Questions of segmentation fidelity, inference latency, robustness to occlusion, and sensitivity to lighting or weather variations are all crucial when translating laboratory benchmarks into operational decision-support tools [4]. Moreover, the relative performance trade-offs between model accuracy and computational footprint directly impact choices about hardware platforms, power consumption, and real-time responsiveness [5].

The purpose of the work is comparative analysis of effectiveness of using MobileNetV3 and ViT neural network models for graphical localization of destroyed buildings remains areas.

To meet this objective, a dedicated software suite was architected and implemented, enabling neural network based analysis of imagery depicting the debris of collapsed structures. The system ingests photographic frames captured by cameras mounted on robotic platforms operating directly within demolition or disaster zones.

The core inputs to the training pipeline consist of two neural network architectures – Vision Transformer (ViT) and MobileNetV3 – and a purpose-built dataset containing labeled images of various ruin elements. During the training phase, each image undergoes preprocessing tailored to the specific requirements of each model: for MobileNetV3, frames are resized to 224×224 pixels, converted into tensor format, and normalized to a [0, 1] range; for ViT, identical resizing and tensor conversion steps are applied, with pixel values scaled to the [–1, 1] interval.

In the subsequent stage, both models are fine-tuned on the prepared dataset. Their performance is then assessed using established accuracy metrics – such as precision, recall, and F1-score – to identify the best-performing parameter configurations, which are persistently stored for later use.

Following model optimization, the next phase of development integrates these trained networks into a user-facing application [6]. This software is then rigorously evaluated in realistic operational settings to confirm its robustness, reliability, and analytical accuracy under conditions that closely mimic actual field deployment.

At deployment, the solution leverages the pre-trained ViT and MobileNetV3 engines to automatically process new images of destroyed building remains, producing concise analytical reports that summarize the detected structural elements and their spatial distribution.

The graph (Fig. 1) shows the ROC curves for the ViT multi-class classification model of construction waste. The ROC curve for each class demonstrates the ratio between the level of false positives and the level of true positives. The area under the curve (AUC) is a metric that characterizes the ability of the model to separate one class from others.

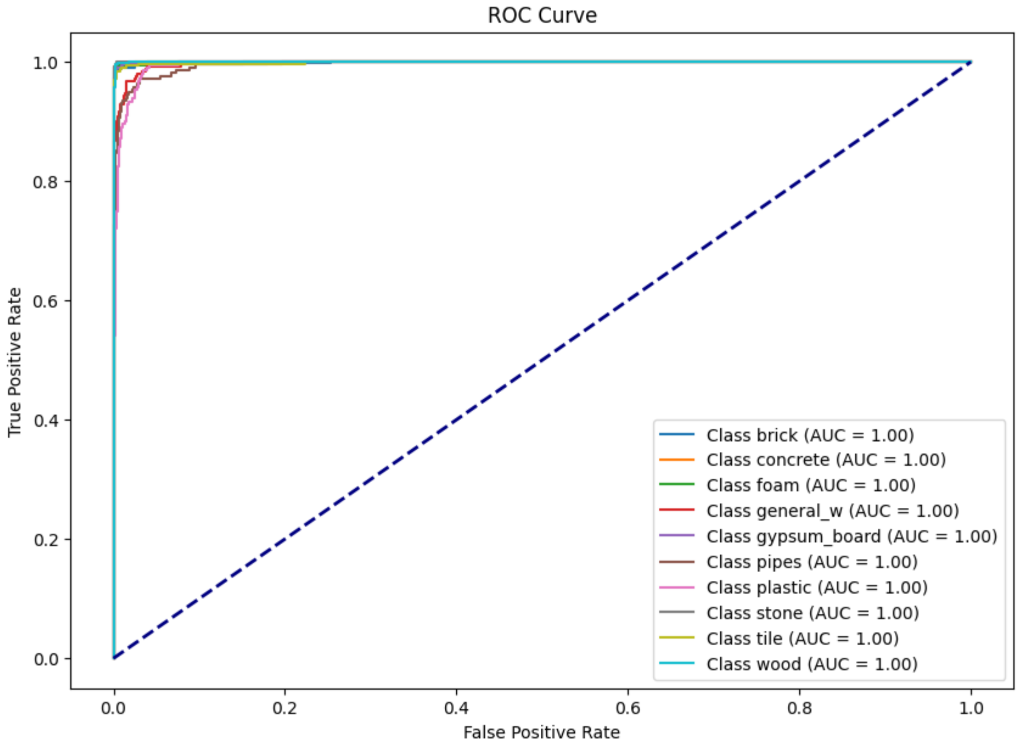

Fig. 1. **ROC curves for the ViT neural network model**

Based on the presented ROC graph (Fig. 2), obtained as a result of the classification of construction waste using the MobileNetV3 neural network, we can draw conclusions about the high discriminative ability of the model. The area under the ROC curve (AUC) indicators approach 1.00 for all classes, which indicates that the model almost accurately separates positive samples from negative ones within each class. The curves are concentrated in the area of high level of true positives and low level of false positives, which is typical for models with high accuracy.

Overall, the results demonstrate the effectiveness of MobileNetV3 as architecture for multi-class object classification tasks in the field of construction debris image processing. For the final assessment of the suitability of the model for practical use, it is necessary to additionally take into account Precision-Recall metrics, error analysis and stability on independent test samples.

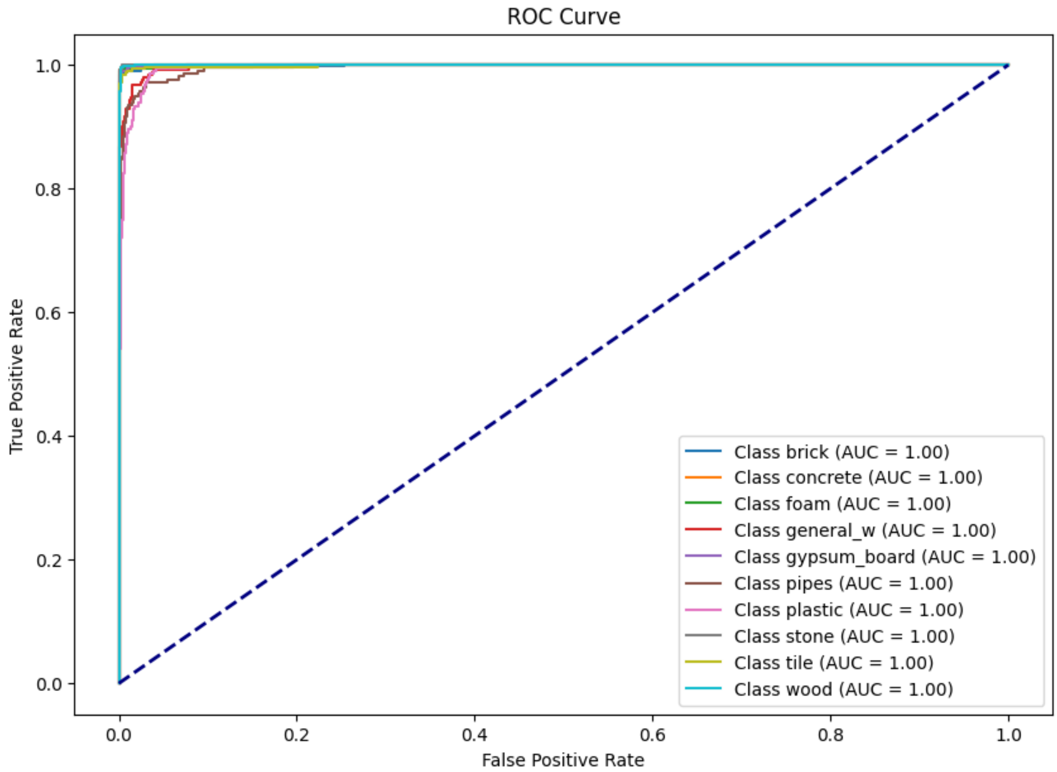

Fig. 2. **ROC curves for the MobileNetV3 neural network model**

As a result of the conducted research, software was developed for neural network analysis of photo data of the remains of destroyed buildings, obtained from robotic equipment. A full cycle of data processing has been implemented – from image preparation and model training to the creation of an application for the automated classification of fragments of building materials.

References:

1. Demetriou, D., Christodoulou, S., Komodromos, A., & others. (2023). Real-time construction demolition waste detection using state-of-the-art deep learning methods: Single–stage vs two-stage detectors. Waste Management, 167, 194–203.
2. Pokhytun, A., Mazurets, O., Molchanova, M., & Tyschenko, O. (2024). Method for Neural Network Detecting Changed Images of People's Faces Using CNN. New Horizons in Scientific Research: Challenges and Solutions. Proceedings of the 1st International Scientific and Practical Conference, 35-40.
3. Mazurets O. V., Klimenko V. I., Molchanova M. O., & Sultanov A. V. (2024). Object-Oriented Intelligent System for Neural Network Detection of Sugar Crystallization Zones. Global Science: Prospects and Innovations. Proceedings of the 10th International scientific and practical conference. Cognum Publishing House. Liverpool, United Kingdom, 198-207.
4. Mazurets, O., Zalutska, O., Tyschenko, O., & Bohdanova, A. (2024). An Approach to Using MobileNet CNN-model for Gesture Recognition. Proceedings of XXIII International Scientific and Practical Conference «Problems of Science and Technology: the Search for Innovative Solutions», 59-64.
5. Zharnovskyi, O., Mazurets, O., & Sobko, O. (2024). Approach to Identification of Artificial Intelligence-Generated People Images by Means of Machine Learning. Key Aspects of the Development of Scientific Research in Modern Conditions. Proceedings of the XLV International scientific and practical conference, 69-73.
6. Kharysh I., Sobko O., & Mazurets O. (2024). Designing CNN Neural Network Model for Detecting Fractures of Lower Extremities by X-ray Images. The Impact of Scientific Research on the Development of the Modern World. Proceedings of the XLIV International scientific and practical conference. Dubrovnik, Croatia, 91-96.