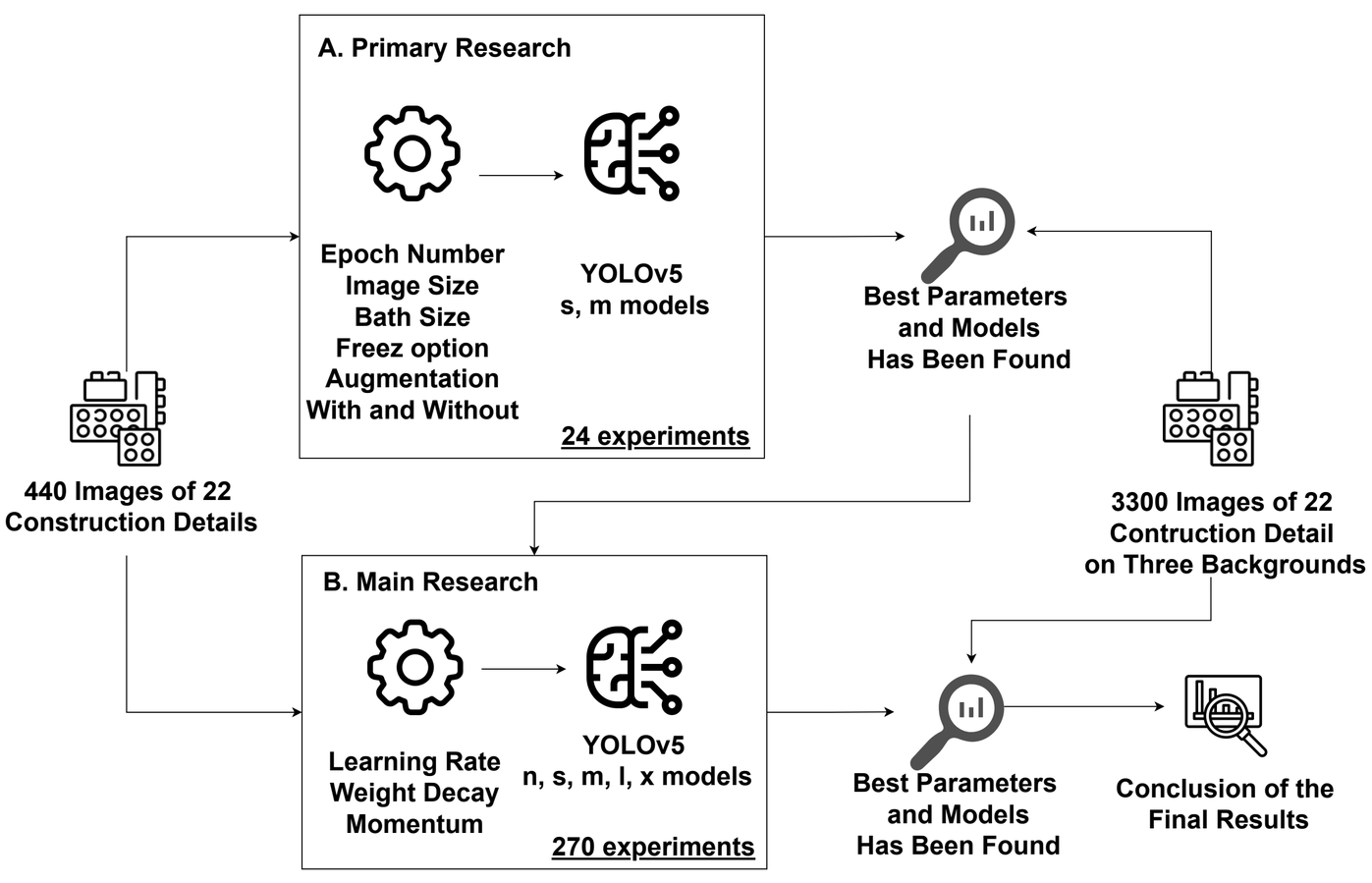
VILNIUS TECH
Faculty of Fundamental Sciences

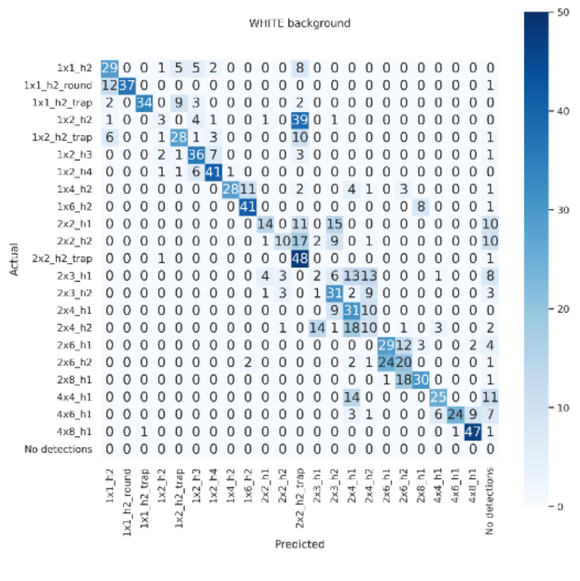
EFFICIENCY OF YOLOV5 MODELS IN THE DETECTION OF CONSTRUCTION DETAILS

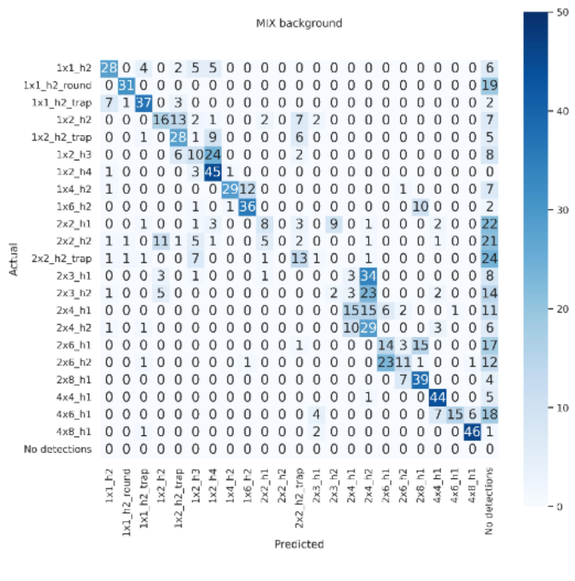
Tautvydas Kvietkauskas, Ernest Pavlov, Pavel Stefanovič

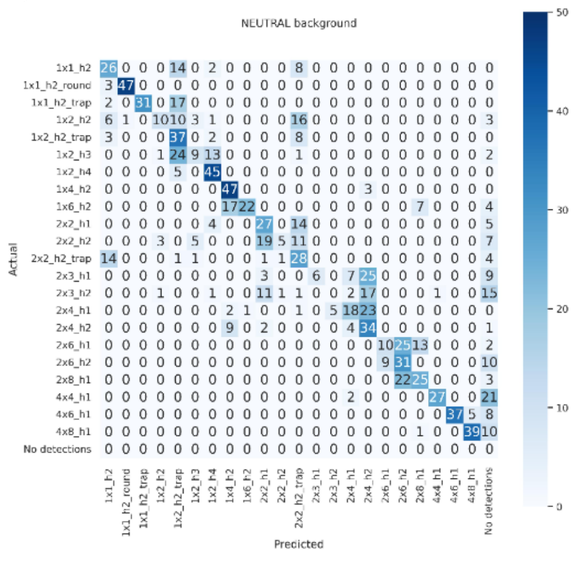
Abstract. Object detection today is widely used in different areas, for example, medicine, industry, business, and even everyday solutions. New models of object detection are constantly developed and the old models are improved by adding some new features or changing the architecture of models. One of the most used object detection models in scientific research is YOLO. In recent years, some new versions of YOLO have been proposed, but they are not fully investigated and lack scientific research. The most stable version of the YOLO group algorithm is the YOLOv5. In this research, the newly collected construction detail dataset has been prepared. Images from different angles of construction details have been taken, and the dataset has been labelled. The dataset consists of 22 construction details. The experimental investigation has been performed using five different models of YOLOv5: n, s, m, l, x. During the experimental investigation, various parameters have been used to find out the influence of the parameters on the final detection results. The models have been tested on three different backgrounds: white, neutral, and mixed. The results of the experimental investigation are promising and, in the future, the models can be used in construction recommendation models.



- The new dataset of construction details has been collected (22 classes).
- The training dataset has been prepared consisting of 440 images (22 construction details on a white background; each class has 20 images).
- The test dataset has been prepared consisting of 3300 images (22 construction details on three different backgrounds; each class has 50 images).
- The five models of YOLOv5 have been investigated: n, s, m, l, x.
- The primary research has been performed to determine the influence of the epochs, batch size, image size, freeze option, augmentation parameters on the results of YOLOv5 models.
- The best parameters of primary research have been found: epoch size - 300, batch size - 32, image size - 320 × 320, freeze option - 10, augmentation on.
- According to the related works, the influence of three parameters has been analyzed: learning rate, momentum, and weight decay.
- Additional research has been performed, to train all models using the higher number of epochs (+100, +300, +600).







Conclusion. The influence of various hyperparameters on model accuracy, including augmentation, learning rate, weight decay, and momentum has been investigated. The most accurate model was YOLOv5l, which recognized 1,655 of 3,300 details across 22 classes. The model was trained with these parameters: an epoch size equal to 600; batch size equal to 32; chosen image size 320 × 320; a freeze option equal to 10; and data augmentation was used. The freeze option reduced training times by approximately two times and construction detail recognition accuracy by approximately 1.5 times. Additionally, the model learning rate was set at 0.001, momentum at 0.95, and weight decay at 0.0001. Future research must be performed to investigate the influence of batch normalization and dropout methods. The research has shown that to avoid overfitting of the training models, in some cases, it may be helpful to increase the batch size. To increase the classification accuracy results, recognition should be done in two steps: by using one YOLOv5 model to recognize construction details in the image and another to create a classification model for each detail separately.