Oleksandr Mazurets
Ph.D. in Engineering Science, Associate Professor of Computer Science Department
Oleksandr Ovcharuk
Postgraduate student
Roman Vit
Postgraduate student
Khmelnytskyi National University, Ukraine

PRACTICE INVESTIGATION OF NEURAL NETWORK DETECTING OF CONSTRUCTION WASTE BY PHOTOS

Abstract. This paper presents a practice-oriented investigation of neural architectures for visual recognition of construction and demolition (C&D) waste in realistic conditions. We study a two-stage pipeline that first localizes candidate fragments on a scene and then assigns material-specific labels, emphasizing data curation, restrained augmentation, and evaluation protocols aligned with production constraints. Results indicate that coupling a single-shot detector with per-class residual classifiers improves recall on visually confusable materials while preserving precision, offering a pragmatic path toward reliable sorting-line deployment.

Rapid urban growth and post-reconstruction have intensified C&D waste streams and exposed the limits of manual inspection and rule-based vision under dirt, weathering, highlights, and occlusion. Deep networks are robust when data mirror operating conditions, but field viability depends more on cross-scene stability and the ability to separate near-identical textures (e.g., light concrete vs. limestone, weathered wood vs. fiberboard) [1]. This study adopts a practice-minded design with minimal assumptions about image capture, conservative class-balance expectations, and metrics tuned to costly error modes.

The recognition system operates in two stages: a single-shot detector first localizes fragments and suggests coarse material types, while a set of one-vs-rest residual classifiers refines these predictions. The final label combines the most confident binary output with the detector's hypothesis to resolve ambiguities, reducing cross-class interference and enhancing separability for confusable categories. The dataset merged open C&D sources with real-world fragment photos, ensuring diversity in background, lighting, contamination, and occlusion to match deployment conditions. Augmentation remained moderate – limited geometric and photometric changes – to mimic camera variability without distorting material identity and to balance minority classes [2].

Implementation favored reproducibility and portability. The detector used moderate input resolution to balance thin-fragment sensitivity with throughput; residual classifiers consumed aspect-ratio-preserving crops normalized to architecture statistics. Early stopping and light regularization limited overfitting on scarce classes. At inference, a calibrated consensus accepted a binary label when its posterior exceeded a margin and did not contradict the detector; ties defaulted to the detector to preserve spatial context.

The study evaluated the accuracy gains from adding specialized models, the effect of augmentation breadth, and robustness under distribution shifts. The base detector showed high precision but missed thin or occluded fragments. Incorporating binary specialists improved recall without increasing false positives, indicating sensitivity to micro-textures and edge geometry. Moderate augmentation – rotations, slight affine shifts, and controlled color variations – stabilized minority classes without raising calibration error, confirming that limited data enrichment is effective..

Robustness was tested on hold-out scenes from different cameras and backgrounds. Although absolute scores dipped with illumination and scale shifts, the two-stage pipeline preserved inter-class ordering and a favorable precision–recall balance versus a monolithic baseline. Errors clustered around boundary crops truncating discriminative features, heavy soiling altering apparent albedo, and boxes containing mixed materials. Arbitration mitigated the first two but not the third, indicating value in instance segmentation or multi-label recognition when warranted.

System-wise, the hybrid design balances accuracy and maintainability. Detectors efficiently propose spatial hypotheses; binary specialists trained with hard negatives add discriminative sharpness where needed. Modularity simplifies evolution: adding or redefining a class requires only the relevant specialist and threshold updates, not full re-training. Data quality remains the dominant lever: coverage across lighting, contamination, and scale governs generalization more than incremental architectural tweaks. Continuous dataset stewardship – sampling from production, rapid relabeling of failure clusters, cautious pseudo-labels – and lightweight domain adaptation (batch-norm recalibration or shallow adapters) provide predictable improvements without disruptive retraining.

In sum, a two-stage neural pipeline for photo-based C&D waste recognition proves viable in realistic settings. Pairing scene-level detection with per-class residual refinement and a simple arbitration rule yields recall gains on confusable materials while preserving precision, reducing misroutes and stabilizing downstream processing. Future work will emphasize multi-label handling of mixed fragments, instance-level segmentation to reduce boundary truncation, and light-touch domain adaptation for rapid cross-site transfer, supporting circular-economy operations through more reliable visual triage.

References:

1. M. Molchanova, V. Didur, O. Mazurets, O. Sobko, O. Zakharkevich, Method for Construction and Demolition Waste Classification Using Two-Factor Neural Network Image Analysis, in: CEUR Workshop Proceedings, 3970, 2025, pp. 168–182.

2. O. Mushtyn, O. Sobko, M. Molchanova, O. Mazurets, Convolutional Neural Network Architecture for Image-Based Architectural Style Recognition, in: Evolving Science: Theories, Discoveries and Practical Outcomes, Proceedings of the 4th International Scientific and Practical Conference, Zurich, Switzerland, 2025, pp. 130–143.