Objective Video Quality Assessment Method for Object Recognition Tasks

Mikołaj Leszczuk, Lucjan Janowski, Jakub Nawała and Atanas Boev

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
AGH University of Krakow

03.07.2024

Introduction to Objective Video Quality Assessment Method for Object Recognition Tasks

- Video quality assessment is crucial for object recognition tasks in fields like surveillance and telemedicine.
- Traditional methods focus on human perception, not machine vision needs.
- The need for objective video quality assessment tailored to machine vision for better accuracy and reliability.
- Addressing real-world challenges such as occlusion, poor lighting, and varying weather conditions.

Research Gap in Video Quality Assessment

- Existing assessments overlook fluctuating lighting, motion blur, and obstructions.
- Need for machine vision-specific assessments.
- Existing methods fail to meet the requirements of Target Recognition Videos (TRVs).
- Challenges in video quality include under-/over-exposure, blurring, noise, and compression artifacts.
- The gap between human visual quality assessment and machine vision needs.
- Importance of developing robust methods to assess video quality for object recognition.

Research Questions and Hypotheses

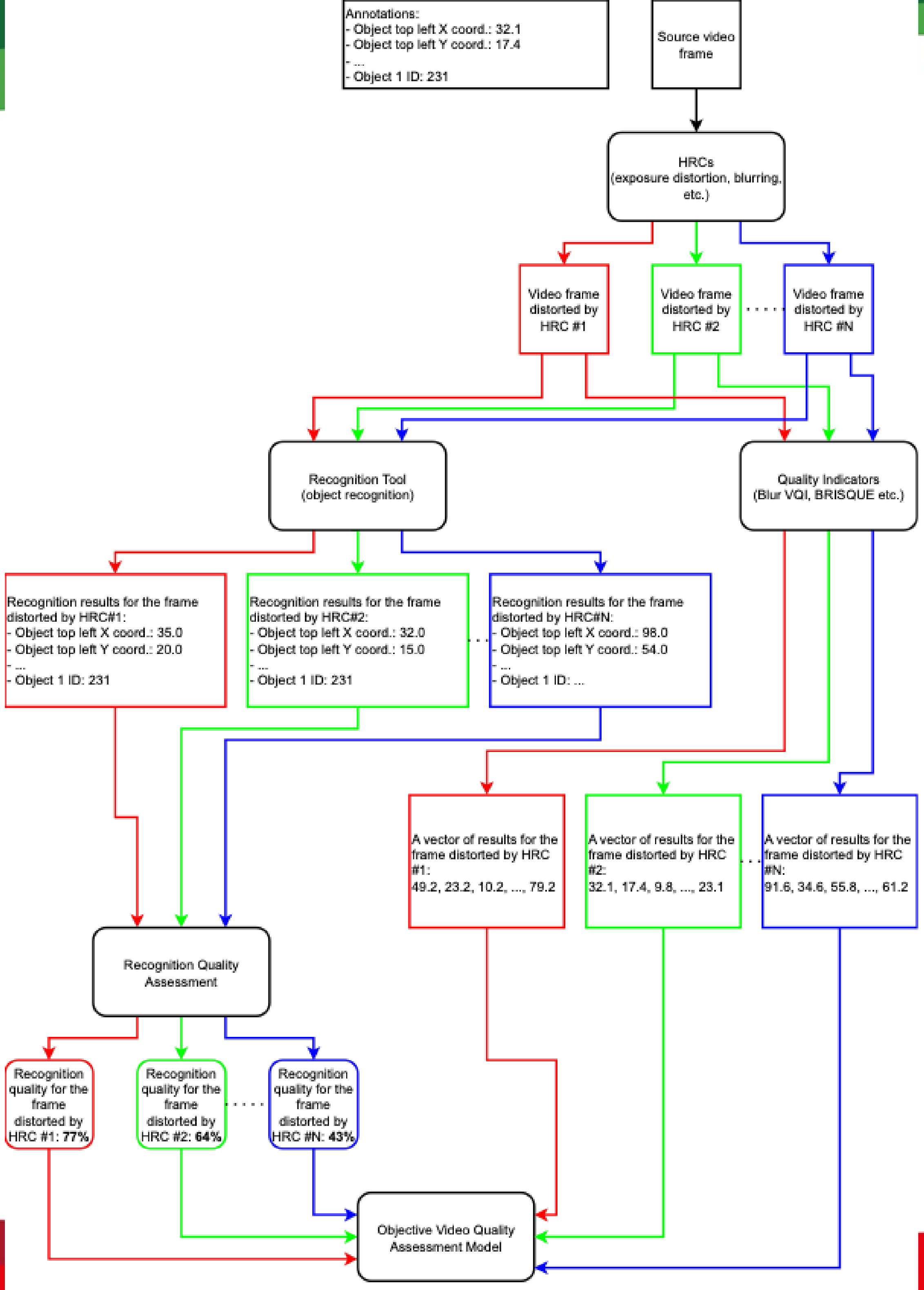
- How can video quality assessments be tailored for TRVs?
- What role do advanced machine learning algorithms play in enhancing feature detection for TRVs?
- Hypothesis: Integrating machine learning improves the precision and reliability of TRV quality assessments.
- Addressing fluctuating lighting and motion blur to enhance object recognition accuracy.
- Evaluating the impact of environmental factors on video quality and object recognition.
- Developing metrics that are more relevant to machine vision tasks.

Methodology

- Use of advanced machine learning models and custom Video Quality Indicators (VQIs).
- Dataset handles real-world challenges like occlusion and poor lighting.
- Flowchart illustrating the interactions among Recognition Experiment, Quality Experiment, and Objective Video Quality Assessment Model.
- Implementation steps include data collection, preprocessing, and model training.
- Leveraging large datasets to ensure comprehensive assessment.

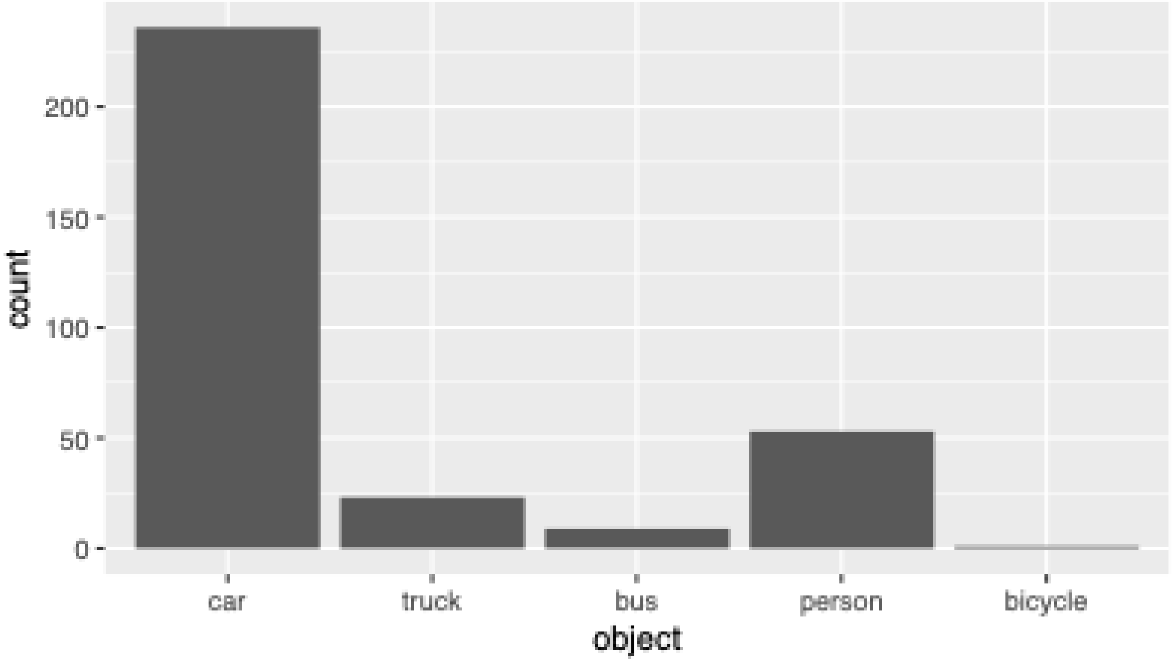
Methodology (continued)

- Video data loading, distortion application, recognition processing, and quality assessment steps.
- Evaluation of different distortions and their impact on object recognition performance.
- Use of statistical and machine learning techniques to analyse results.
- Steps to ensure the reproducibility and reliability of the assessment method.
- Detailed workflow for systematic video quality assessment.
- Data augmentation techniques to create diverse and challenging scenarios for evaluation.
- Statistical analysis methods to validate the effectiveness of the proposed VQIs.



Source Reference Circuits (SRCs)

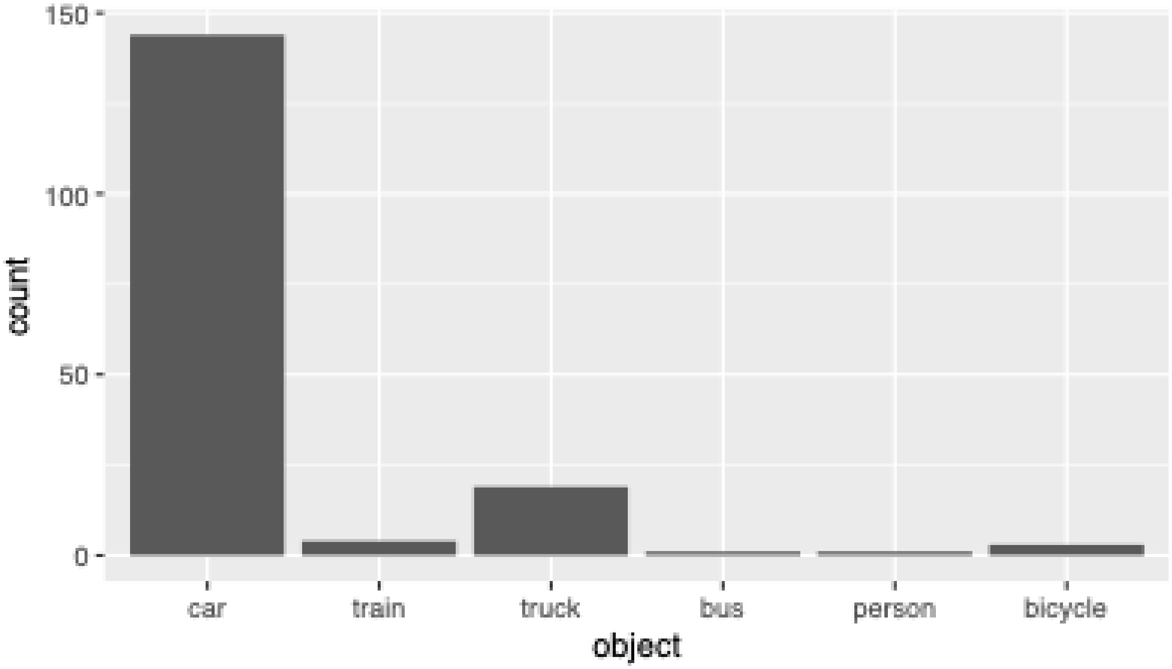
- Integration of nuScenes and KITTI datasets for diverse video sequences.
- Selection criteria based on processing time and object coverage.
- Statistical analysis of selected video frames from the nuScenes dataset.
- Ensuring a wide variety of urban and rural scenes for robust testing.
- Addressing diverse environmental conditions to validate the assessment method.



Source Reference Circuits (SRCs) (continued)

- Example frame from the nuScenes dataset depicting typical urban traffic conditions.
- Next slide:
  - Statistical analysis of selected video frames from the KITTI dataset.
  - Example frame from the KITTI Vision Benchmark Suite dataset depicting typical urban traffic conditions.
  - Addressing challenges such as varying lighting conditions and motion blur.

Source Reference Circuits (SRCs) (continued)



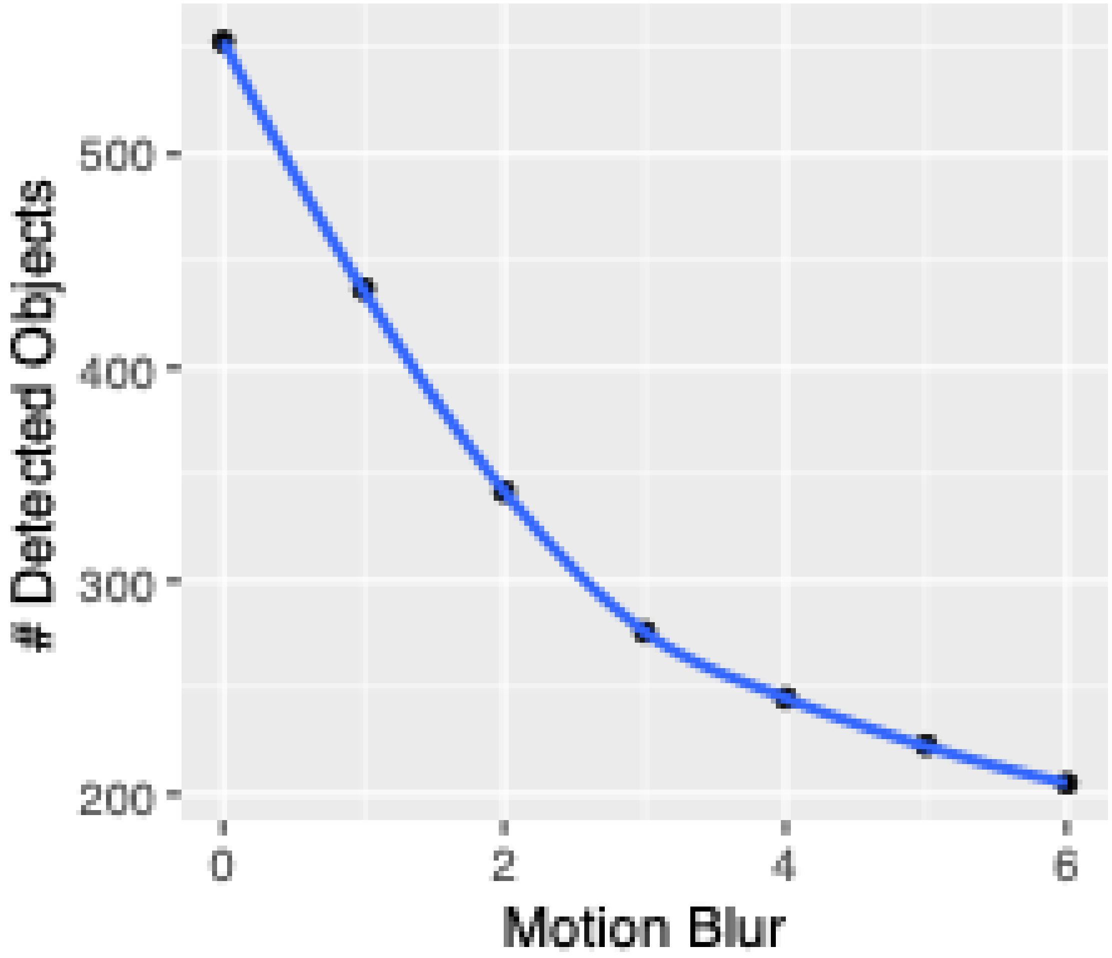
Source Reference Circuits (SRCs) (continued)

Hypothetical Reference Circuits (HRCs)

- HRCs include under-/over-exposure, defocus, Gaussian noise, motion blur, and JPEG compression.
- Tools used: FFmpeg and ImageMagick.
- Thresholds for distortions are set to identify the point at which recognition fails.
- Computational performance evaluated under maximum load.
- Detailed analysis of how different distortions impact object recognition accuracy.
- Use of synthetic distortions to test the limits of object recognition systems.

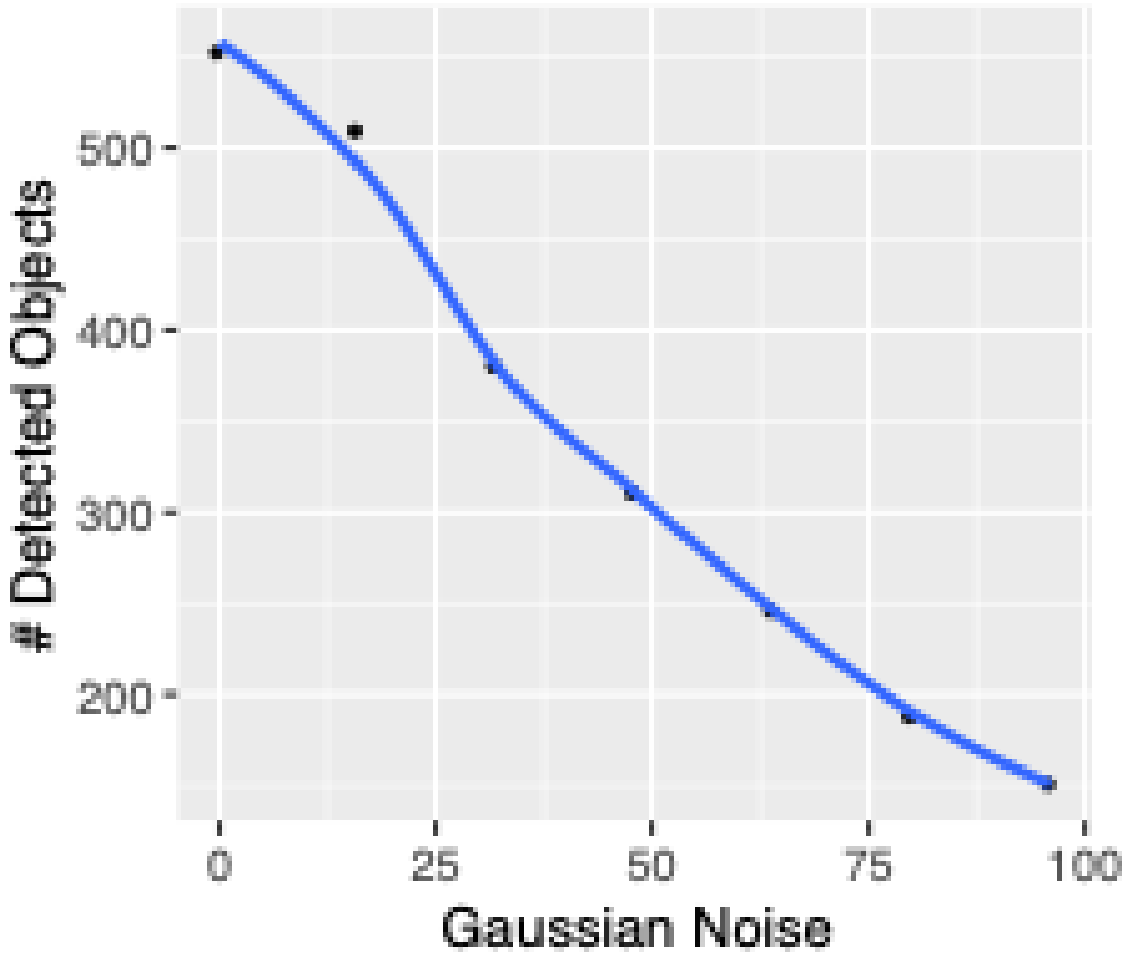
Recognition Experiment

- Object detection using YOLOv3 trained on the COCO database.
- Results stored in structured JSON format for analysis.
- Comparison of recognition performance under various distortions.
- Analysis of the impact of motion blur on object recognition performance.
- Detailed evaluation of object detection accuracy under different conditions.
- Examining the robustness of YOLOv3 in diverse environments.
- Insights into the model's performance variations due to quality degradation.

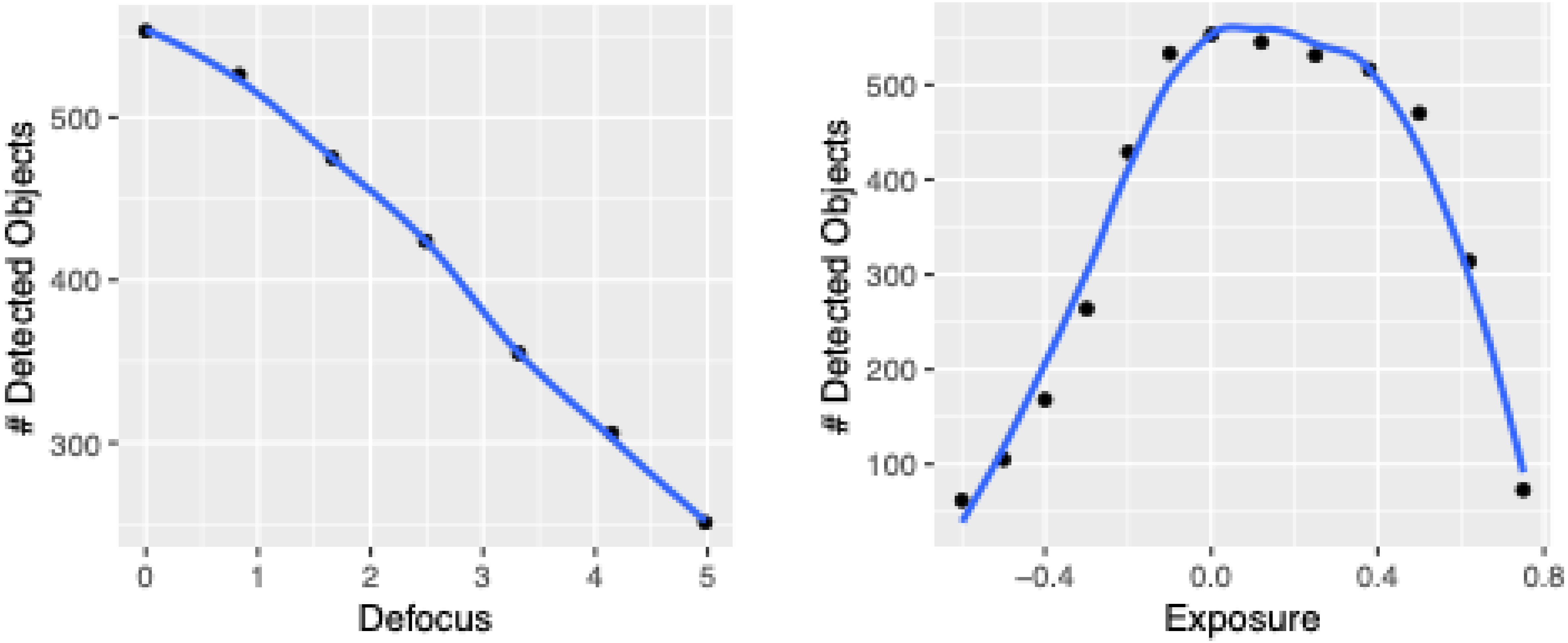


Recognition Experiment (continued)

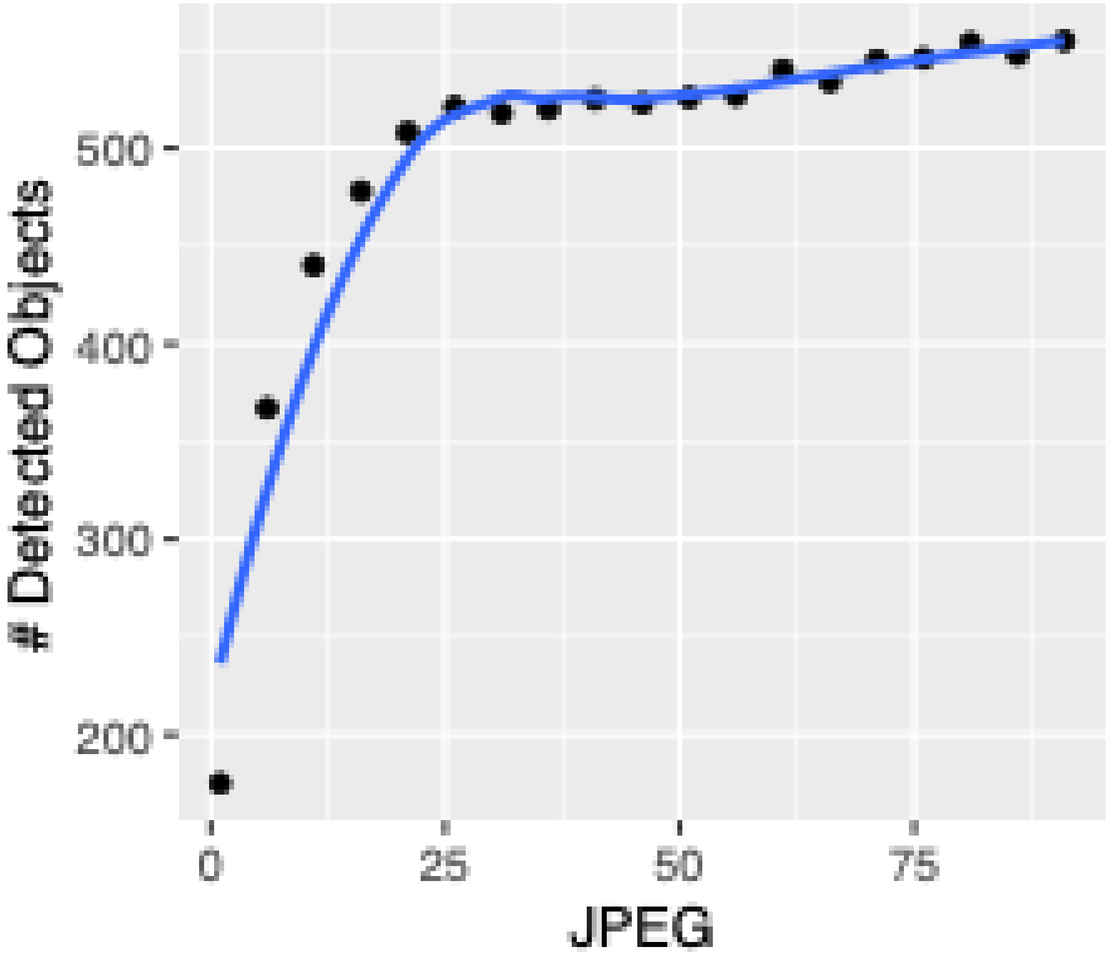
- Analysis of the influence of different distortions on the number of objects detected.
- Results demonstrate the robustness of the YOLOv3 model under varying conditions.
- Importance of maintaining video quality for accurate object detection.
- Examination of the effects of Gaussian noise on recognition performance.

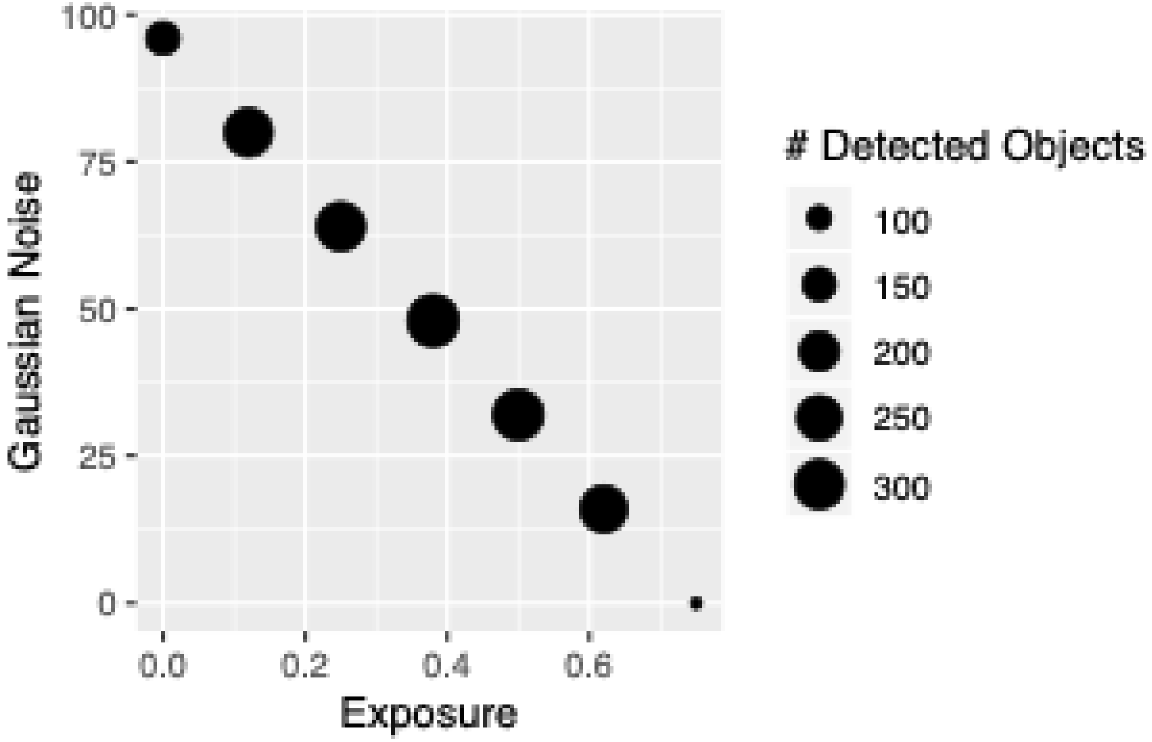


Recognition Experiment (continued)

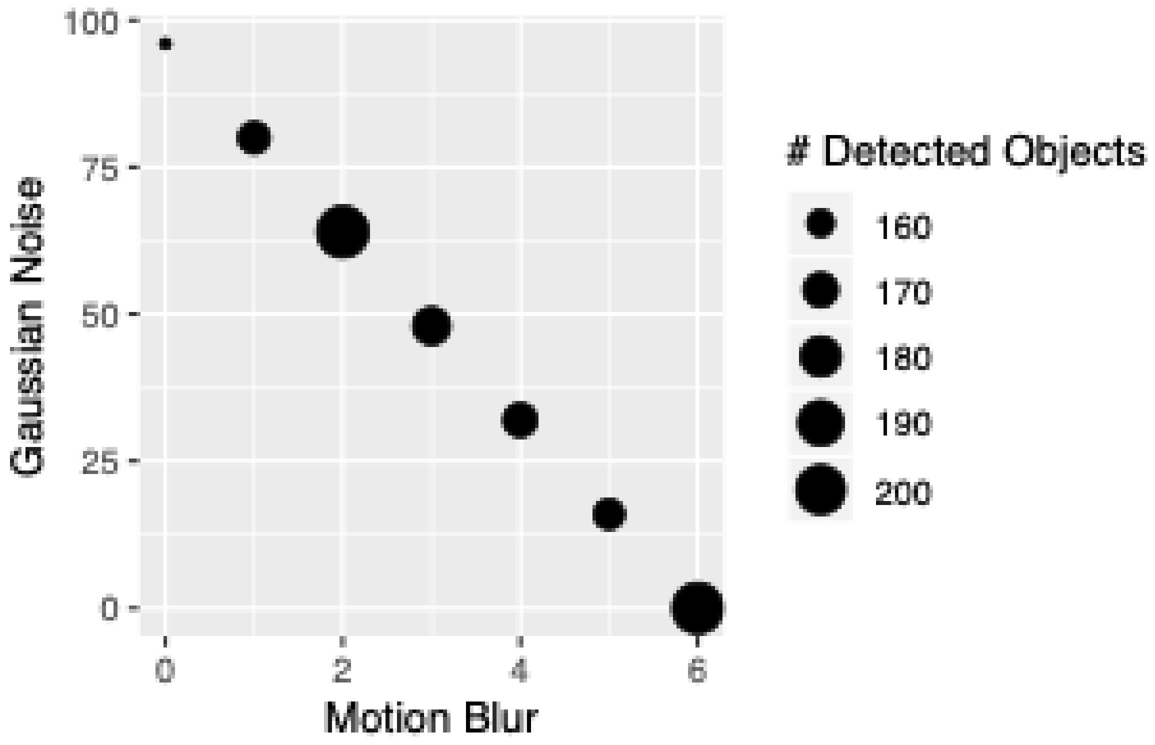


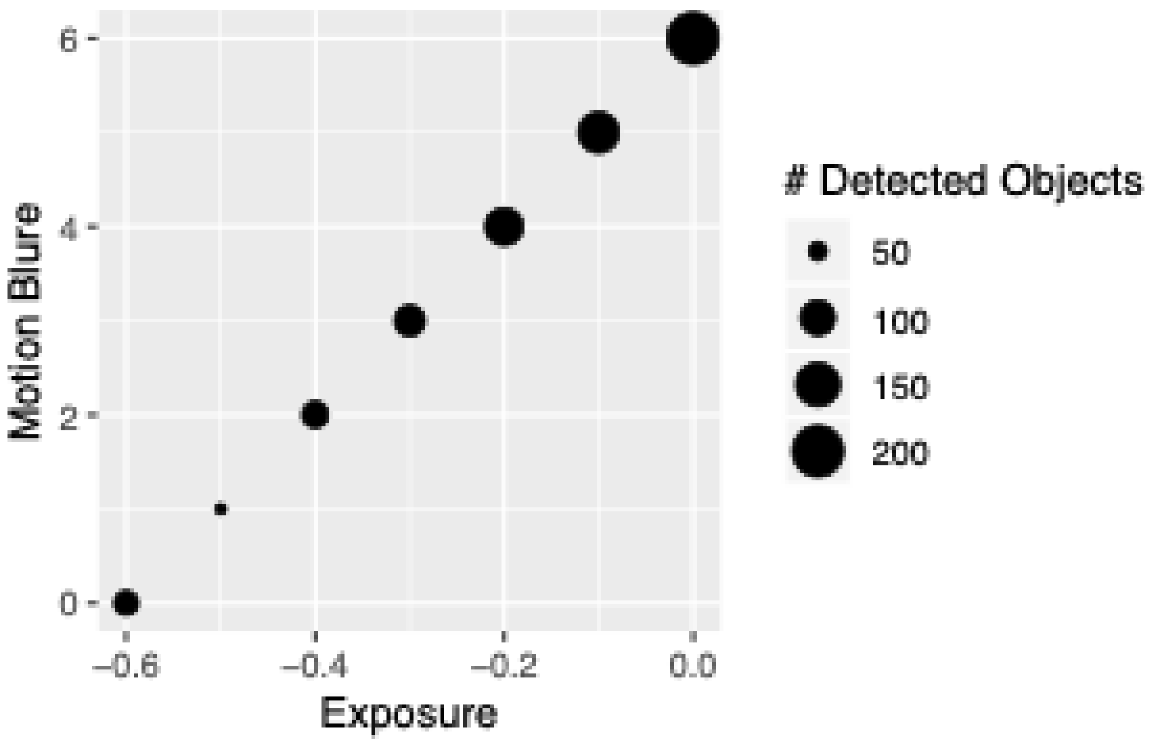
Recognition Experiment (continued)





Recognition Experiment (continued)



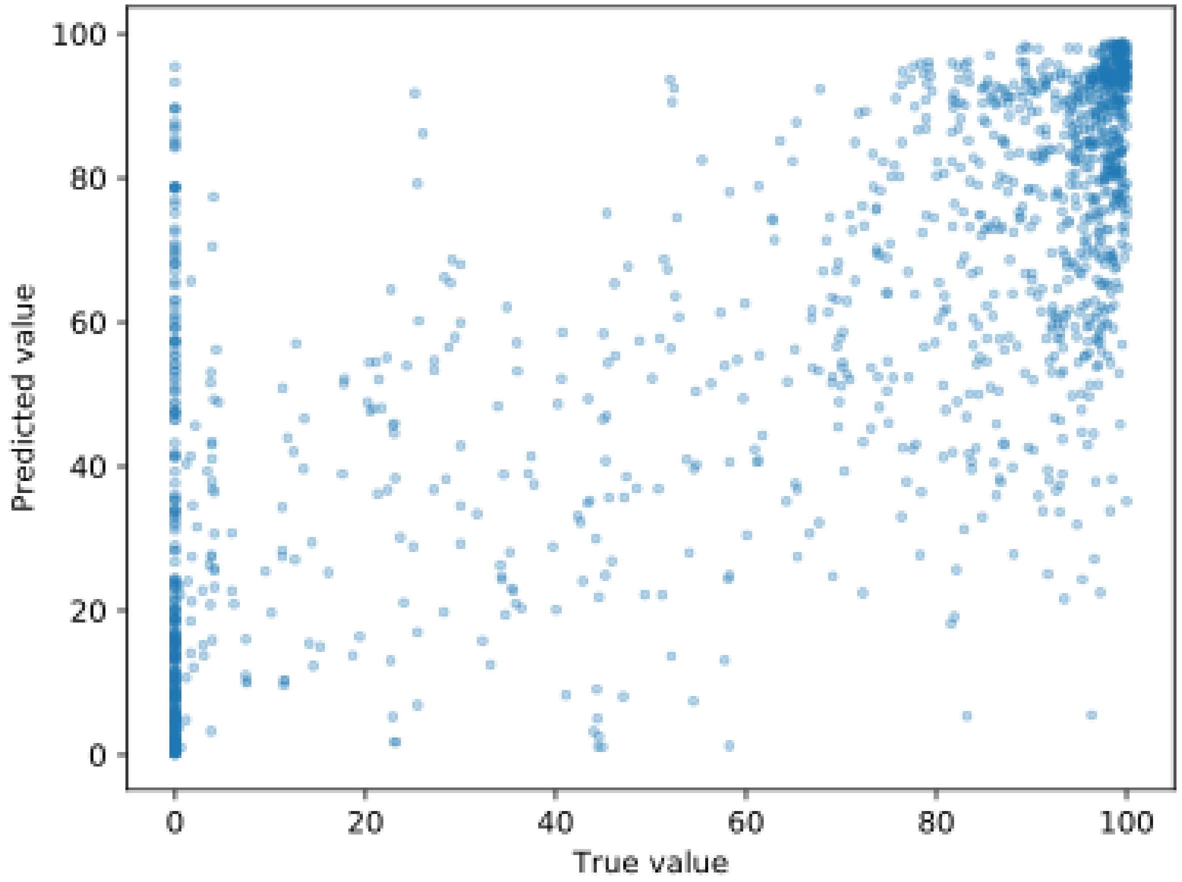


Quality Experiment

- Evaluation of various Video Quality Indicators (VQIs).
- Focus on individual video frames and execution times.
- Categories: 'All Metrics' and 'Our (AGH) Metrics'.
- Detailed methodologies and computational frameworks are provided in cited papers.
- Analysis includes metrics like Blockiness, Blur, Contrast, Exposure, Noise, and Spatial Activity.
- Comprehensive approach ensures robust quality assessments tailored to object recognition tasks.
- Comparing different VQIs to identify the most effective for machine vision.

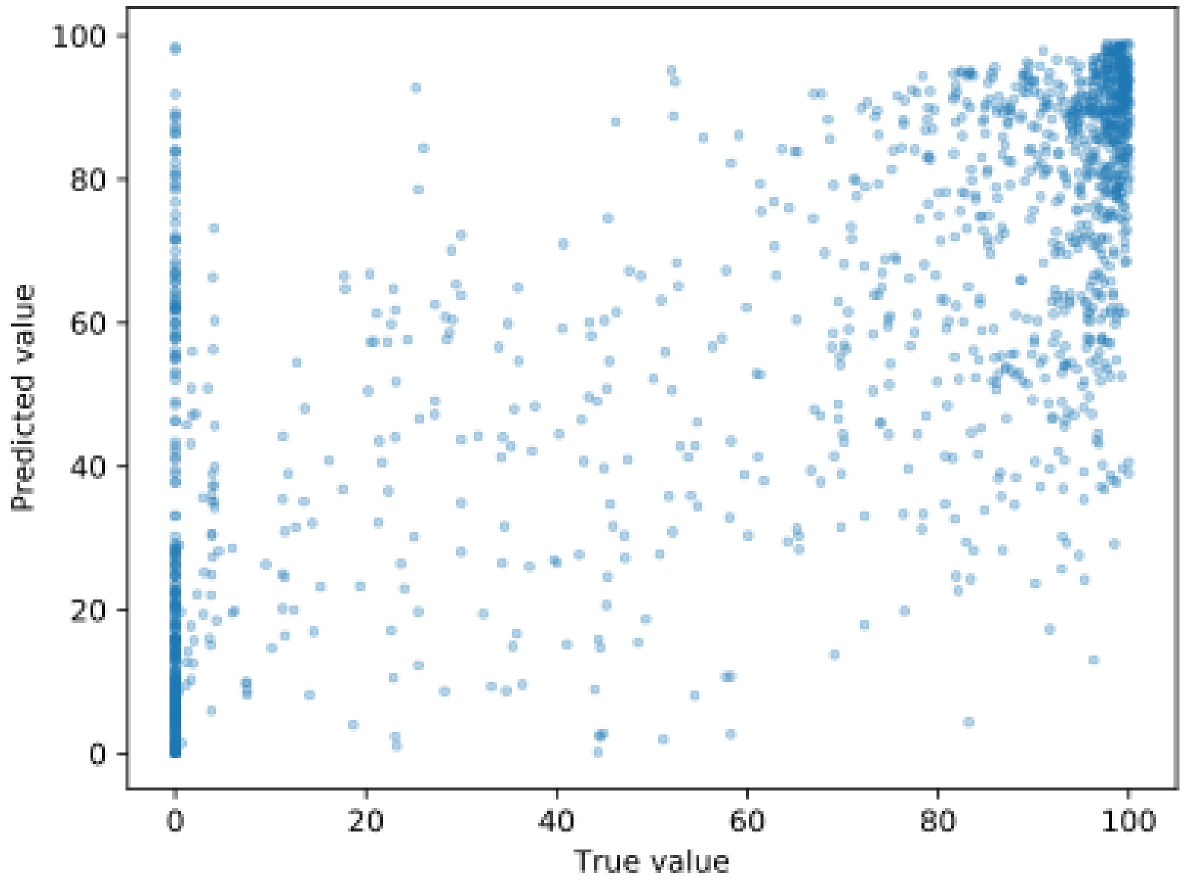
Results

- Random forest regressor model achieved MSE of 672.4 and correlation of 0.77.
- Scatter plots show correlation between predicted and actual values.
- Performance comparison between different VQIs.
- Detailed analysis of the effectiveness of the proposed methodology.



Results (continued)

- Using only AGH metrics, MSE was 722.1 with a correlation of 0.75.
- Models effectively classify recognition performance under varying conditions.
- Results highlight the robustness of the proposed methodology.
- Future work to refine models and expand datasets for improved accuracy.



Conclusion

- Significant advancement in video quality assessment for TRVs.
- Future work to expand dataset and integrate diverse machine learning algorithms.
- Plans for collaboration with industrial and academic partners.
- This research aims to enhance operational effectiveness in surveillance and telemedicine.
- Addressing limitations such as computational resource constraints and dataset diversity.
- Potential to revolutionize operational effectiveness in high-stakes environments.
- Broader applicability expected in varied real-world scenarios.

Reference
Questions and Answers

- Leszczuk, M., Janowski, L., Nawała, J., & Boev, A. (2024). Objective Video Quality Assessment Method for Object Recognition Tasks. Electronics, 13(9), 1750. https://s.agh.edu.pl/gY1dU
- Open floor for questions from the audience.