Automated Detection and Characterization of Aluminum Agglomerates in Solid Propellant Combustion

Omri Troytza *, Ilan Halperin † and Anton Nezer‡
Afeka Academic College of Engineering, Tel-Aviv, Israel, 6998812

Dima Bykhovsky §
Shamoon College of Engineering, Beer-Sheva, Israel, 8410802

Oshrit Hoffer ¶ and Yinon Yavor ‖, **
Afeka Academic College of Engineering, Tel-Aviv, Israel, 6998812

This is the author's accepted manuscript of an article published in the *Journal of Propulsion and Power*.
Troytza, O., Halperin, I., Nezer, A., Bykhovsky, D., Hoffer, O., and Yavor, Y., "Automated Detection and Characterization of Aluminum Agglomerates in Solid Propellant Combustion," *Journal of Propulsion and Power*, 2026, Articles in Advance.
`https://doi.org/10.2514/1.B40472`
Published version in the AIAA electronic library: `https://arc.aiaa.org/doi/10.2514/1.B40472`

Aluminum agglomeration in solid rocket propellants has a significant impact on two-phase flow losses and combustion efficiency. Characterizing these agglomerates typically relies on high-speed photography, where manual analysis is labor-intensive and prone to subjectivity, while traditional image processing algorithms struggle with noise and motion blur. This study presents an automated pipeline for the detection of agglomerates using a Mask R-CNN deep learning architecture with a ResNet-50 backbone. The model was trained and evaluated on a dataset of 166 high-speed images containing over 8,100 manually annotated instances, captured under varying pressure conditions (2 MPa to 4 MPa). A shortest-chord diameter estimator is applied uniformly to all segmentation masks to reduce the systematic inflation that area- and perimeter-based estimators exhibit on motion-blurred masks. Detection performance reaches an F_1-score of 0.82, and population-level agreement with the manual baseline yields 1% difference in terms of D_{32} (Sauter mean diameter) and 3% difference in terms of D_{43} (volume weighted diameter). Our results demonstrate strong alignment with manual annotations, establishing deep learning as a highly scalable and effective alternative to manual analysis in this imaging regime.

*Undergraduate student, School of Electrical Engineering, 38 Mivtza Kadesh
†Undergraduate student, School of Electrical Engineering, 38 Mivtza Kadesh
‡Undergraduate student, School of Mechanical Engineering, 38 Mivtza Kadesh
§Senior Lecturer, Department of Electrical and Electronics Engineering, Bialik 56
¶Senior Lecturer, School of Electrical Engineering, 38 Mivtza Kadesh
‖Senior Lecturer, School of Mechanical Engineering, 38 Mivtza Kadesh
**Corresponding author, email: YinonY@afeka.ac.il

I. Introduction

Composite solid propellants used in rocket motors generally contain a polymeric binder fuel mixed with an oxidizer in a powder form. In order to enhance the specific impulse (I_{sp}) and density of the propellant, metal powders like aluminum or magnesium (typically 10 µm–40 µm in diameter) are incorporated into the mixture that forms a solid grain which can be readily ignited. However, the use of aluminum in solid propellants introduces the undesired phenomena of agglomeration. This phenomenon involves the accumulation of numerous particles into larger aluminum aggregates [1]. Reducing aluminum agglomeration is of great significance in solid rocket motors because it decreases incomplete combustion and mitigates two-phase flow losses, ultimately resulting in improved motor performance. Specifically, smaller agglomerates reduce the velocity and thermal lag between the condensed-phase particles (aluminum and alumina) and the surrounding gas flow through the nozzle, which minimizes the performance penalty in specific impulse (I_{sp}) and characteristic velocity (c^*).

Developing and validating image-based detection and sizing methods specifically at elevated pressures is motivated by two considerations. First, solid rocket motors operate at chamber pressures of approximately 3 MPa to 10 MPa; agglomerate characterization at atmospheric pressure does not represent actual motor-operation conditions, both in terms of operating pressure and because elevated-pressure experiments use an inert atmosphere, while atmospheric experiments are usually conducted in air, which acts as an additional oxidizer. Second, methodologically, the pressurized combustion environment constrains optical access, increases burn rate, and produces faster particle dynamics, so that a method demonstrated at 2–4 MPa validates its applicability across a broader and more practically relevant regime.

Historically, significant efforts have been dedicated to assessing and predicting the size of agglomerates ejected from the burning surface. High-speed photography has been a primary tool to study agglomeration in metalized solid propellants, as seen in the works of Crump et al. [2], Gany et al. [3], Glotov et al. [4] and DeLuca et al [5]. The agglomeration phenomenon is illustrated in Fig. 1, where the first two frames show the particle accumulation and formation; the third frame shows detachment from the propellant surface; and the final frame shows the protrusion of the oxide cap.

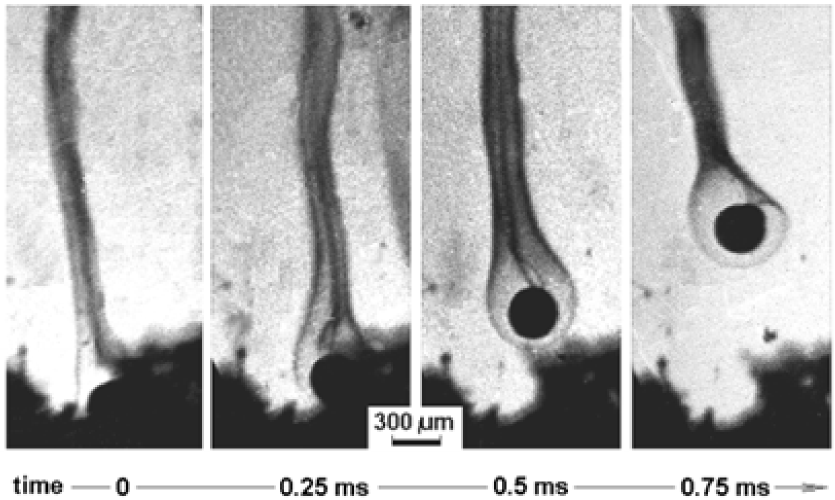


Fig. 1 Agglomeration evolution in composite solid propellant [4].

Theoretical modeling has evolved alongside experimental observations. Beckstead [6] presented a comprehensive overview of agglomeration modeling, distinguishing between the geometric approach (e.g., the Pocket Model [2, 7–9]) and the mechanistic approach (e.g., Gany and Caveny [10]). A more recent model by Yavor et al. [1] merges these two concepts, showing good agreement between model predictions and experimental results [11–13]. Subsequent experimental works [14, 15] focused on coated nano-aluminum particles, using manual analysis of high-speed photography. These studies concluded that while the average agglomerate size was hardly changed, the total volume of ejected agglomerates decreased with the use of coated nano-Al powder. Furthermore, agglomeration was consistently reduced as pressure increased, which was also demonstrated in [16]. Xiao et al. [17] employed a theoretical 3D micro-structural framework that infers aluminum agglomerate size distributions from density-based clustering of aluminum particle positions, achieving similar trends and reasonable agreement with experiments.

Despite these advances in theoretical modeling, experimental agglomerate characterization has traditionally relied heavily on manual annotation. However, significant progress has been made through the integration of automated image processing. Decker et al. [18] combined shadowgraphy with the MSER algorithm to successfully detect agglomerates both before and after their detachment from the burning surface. They also offer a newer curvature-based multiscale method for improved particle identification and measurement, but have yet to test it with aluminum particles in propellants. In a comparative study, Powell et al. [19] evaluated three different methods for counting and measuring aluminized agglomerates. However, the study reported significant variability across methods and lacked validation against manual counts, making accuracy assessments challenging. Furthermore, these tests were limited to atmospheric pressure, necessitating further research at higher pressures.

Recently, deep learning and transfer learning have emerged as robust alternatives for automated image processing.

Architectures such as U-Net have demonstrated high accuracy in droplet segmentation [20]. Advancements by Sportillo et al. [21] have enabled the semantic segmentation of specific droplet components and the estimation of their individual sizes. Other studies have employed ResNet backbones [22, 23] for detection; however, these methods frequently rely on expensive microscopic imaging setups to minimize noise, thereby limiting their general accessibility. In the domain of real-time analysis, a recent study utilized the YOLOv11 instance segmentation model, combined with the ByteTrack algorithm, for frame-to-frame tracking of aluminum agglomerates [24]. Using a dataset of 200 images and 676 labeled agglomerates, this approach achieved an AP50 of 0.94. While this method delivers impressive high-speed performance, the architectural design of YOLO prioritizes real-time inference speed over the segmentation precision required for accurate physical measurement.

The present work addresses several limitations of existing approaches by proposing an automated pipeline for agglomerate detection and measurement in high-speed images of aluminized propellant combustion. We employ a Mask R-CNN architecture with a ResNet-50 backbone to segment agglomerates from images acquired using a relatively simple and accessible high-speed setup. The resulting segmentation masks are processed by a shortest-chord algorithm that computes a diameter estimate D_{sc} from each mask. Adopted uniformly across all masks, D_{sc} serves as the most suitable estimator for this dataset rather than an externally validated measurement of the true diameter of the particles. Importantly, it improves consistency with the manual baseline by mitigating the systematic inflation that area- and perimeter-based estimators exhibit on motion-blurred masks. Using a dataset of 166 high-speed images with 8,109 manually annotated agglomerates, the proposed method achieves strong performance in both detection and sizing, despite the relatively simple imaging configuration. This demonstrates that accurate, automated agglomerate characterization can be achieved without specialized optical instrumentation, thereby enabling more efficient and extensive studies of agglomeration phenomena in solid rocket propellants.

II. Methods

A. Materials

The tested propellant formulation is summarized in Table 1. Mass-wise, it consists of 18% HTPB binder, 2% IPDI cross-linker and 65% ammonium perchlorate (AP) oxidizer. The AP comprises 52% coarse particles (400 µm) and 13% fine particles (90 µm), reflecting a 4:1 coarse-to-fine ratio. Additionally, 15% aluminum powder was used, with 10% of micro-Al (6 µm in diameter) and 5% of nano-Al powder (50 – 70 nm). Nano-aluminum powder is generally added to solid propellants to enhance combustion efficiency and burn rate and to reduce agglomeration. However, the very high viscosity of mixtures containing an increased fraction of nano powders limits the total amount of nano-Al inserted to composite solid propellants [1].

Table 1 Propellant formulations.

Material	Function	Particle size	Mass fraction [%]
HTPB	Fuel binder	-	18
IPDI	Cross-linker	-	2
AP	Oxidizer	400 µm	52
		90 µm	13
Al powder	Metalic fuel	6 µm	10
		50 −70 nm	5

B. Experimental setup

Combustion experiments were conducted in a nitrogen-pressurized chamber with a volume of approximately 1 L and a 25 mm thick glass window to provide optical access for high-speed imaging. The strand was installed on a custom-made Teflon strand holder within the pressure chamber, and nitrogen was injected in order to achieve the required pressures ranging from 2 MPa to 4 MPa. Ignition was initiated using a resistive filament powered by a 10 W supply. The experimental arrangement is illustrated schematically in Fig. 2.

Imaging was performed using a Chronos 2.1 high-speed camera positioned externally to the pressure chamber and focused through the glass window on and above the burning surface. The depth of field, set by the lens aperture and magnification, spans several millimeters along the optical axis, which is sufficient to capture agglomerates immediately above the burning surface. The region-of-interest was sized so that its width covers the strand (5 mm) plus edges and its height reaches 3 mm to 4 mm, depending on the experiment. Because the strand burns in an end-burning configuration (due to epoxy coating on its sides), nearly all ejected particles cross the region-of-interest while remaining within the depth of field. Particles travelling significantly outside the depth of field appear as diffuse, low-contrast blobs and typically lie outside the region-of-interest; in the rare cases in which they fall within it, they are rejected by the confidence threshold of the model. During combustion, no external illumination was used: imaging relies entirely on self-emitted light from the burning propellant surface and the incandescent aluminum agglomerates. Although backlighting or shadowgraphy would suppress flame luminosity and improve raw image quality, the present study deliberately uses this simple self-emission configuration in order to demonstrate the robustness of the deep-learning pipeline without specialized instrumentation. Images and videos were recorded at frame rates between 5000 and 5400 frames per second (fps) with an exposure time of 15 µs and a resolution of 800×480 pixels. The 15 µs exposure balances signal intensity against motion blur: for typical agglomerate velocities of 1–5 m/s, this produces a displacement of 15 µm to 75 µm, corresponding to roughly 1–6 pixels at the imaging resolution, which is explicitly handled by the shortest-chord diameter algorithm described in Section II.E. Spatial calibration was performed prior to each ignition event, using external bright illumination only for this calibration step (not during combustion). An image of the unlit propellant strand was

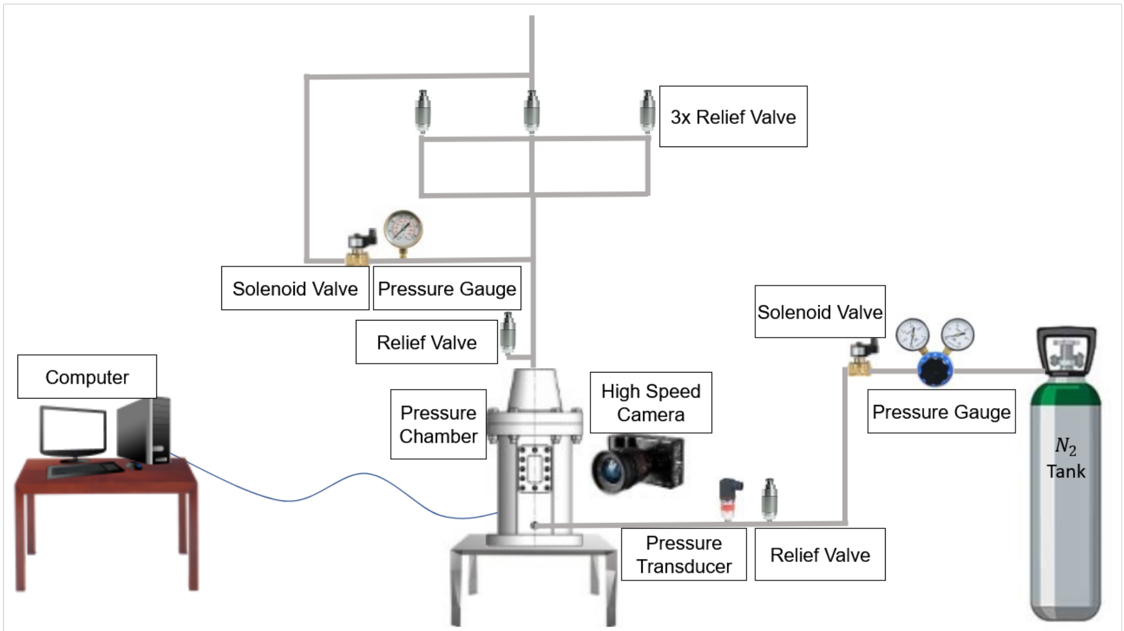
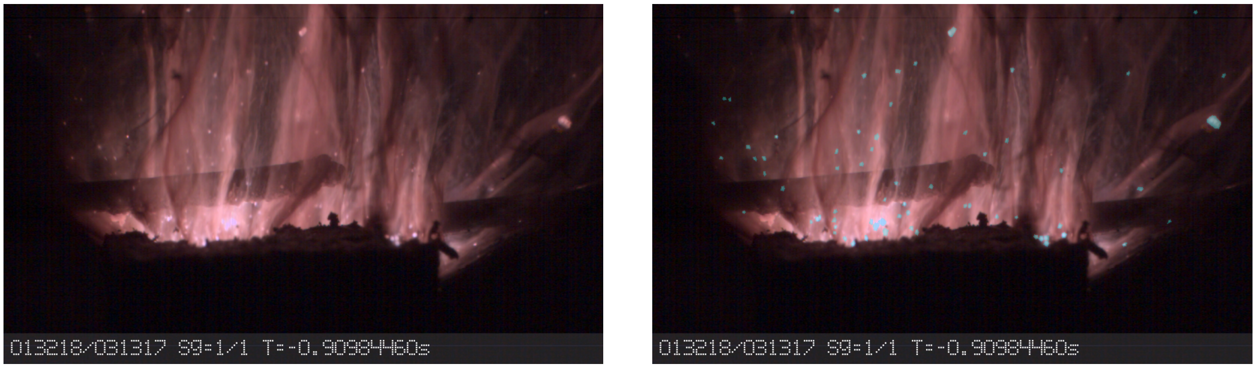


Fig. 2 Schematic illustration of the experimental setup used for high-speed imaging combustion experiments.

captured under this illumination, and the known diameter of the ignition wire (0.5 mm) served as the reference scale. The resulting spatial resolution ranged from 10 µm to 13 µm per pixel, depending on the specific alignment of the strand relative to the focal plane. The analyzed images were taken at combustion pressure of 2–4 MPa, and a typical frame captured during combustion is presented in Fig. 3a, showing the burning surface and particles above it.

C. Dataset

The dataset comprises high-speed images of aluminized propellant combustion, compiled from five firing tests conducted at 2, 3, and 4 MPa. From the high-speed recordings, frames were selected to provide temporal diversity (avoiding consecutive or near-consecutive frames that would contain nearly identical particle populations) and to cover different stages of the combustion process (early, mid, and late burn). Frames with severe obscuration (e.g., dense smoke covering the entire field of view) were excluded. The 166 selected frames therefore represent a diverse sampling of combustion conditions rather than a contiguous sequence from a single test. Because independent physical measurements of the ejected agglomerates are not available in this imaging regime, the ground truth (GT) used throughout this work is defined as the set of manually annotated segmentation masks produced by the authors; all reported diameters on the GT side are obtained by applying the same shortest-chord algorithm used on the model-predicted masks to these manual annotations. The GT therefore represents a human-defined baseline rather than an externally validated physical reference. It was generated via manual annotation using the open-source COCO Annotator framework [25]. Instance-level polygonal segmentation masks were drawn for every visible agglomerate, resulting in a total of 8,109



(a) Original raw image **(b) Ground truth annotation**

Fig. 3 High-speed photography of solid propellant combustion showing the segmentation pipeline. (a) The original raw frame. (b) The image with manual ground truth annotations.

annotated instances drawn across the 166 images. Annotations representing microscopic noise (area ≤ 3 pixels) or unnatural thin linear shapes (e.g., 1×4-pixel lines that correspond to sensor or stitching artifacts rather than real particles) were filtered out prior to training. The pressure-wise distribution is as follows: 5795 annotations from 86 images at 2 MPa, 350 annotations from 39 images at 3 MPa and 1964 annotations from 41 images at 4 MPa. These annotations were serialized in the COCO JSON format, recording the image ID, object category ("agglomerate"), and vertex coordinates for each mask.* An example of an annotated image is presented in Fig. 3b.

D. Deep Learning Model

To validate the performance of the model across the entire dataset, a stratified 10-fold cross-validation was performed. In each fold, 90% of the dataset was used for training and validation, and 10% was held out for testing. The splits were grouped by image, so that all annotations from a given frame were confined to the same fold, preventing information leakage between training and test data. In addition, the folds were stratified by pressure, so that the proportional representation of 2 MPa, 3 MPa, and 4 MPa samples was preserved across every fold, avoiding any bias toward the dominant 2 MPa condition. Temporal correlation was further minimized by sampling discrete, sufficiently spaced frames (see Section II-C) rather than contiguous sequences.

Mask R-CNN with a pretrained ResNet-50-FPN-V2 backbone was selected for its balance of segmentation accuracy and inference speed. As part of a transfer learning approach, the pretrained model was fine-tuned by replacing the original mask prediction head to accommodate two classes: the background and the agglomerates. The remaining layers, including the backbone and feature extractor, retained their pretrained weights to leverage previously learned feature hierarchies. To adapt the architecture to the characteristics of agglomerate imagery, the Region Proposal Network anchors were customized down to 4×4 pixels and augmented with a broader range of aspect ratios, the Region-of-Interest pooling sampling ratio was increased from 2 to 4, the input resolution was raised to up to 2048 pixels,

*The dataset is available online at `https://github.com/bykhov/agglomerate_dataset`.

and the internal per-image detection limit was raised to 3000 during training to accommodate dense frames. Data augmentation included horizontal and vertical flips, spatial rotations of up to ±90°, mild scaling, and brightness/contrast jitter. Hyperparameter tuning was performed through a series of experiments in which learning rates and epoch counts were systematically varied. The optimal configuration was achieved using the AdamW optimizer with an initial learning rate of 5×10^{-3}, combined with weight decay and early stopping mechanisms. Training was conducted on a single NVIDIA RTX 3060 Ti GPU over 100 epochs with a batch size of 2 images.

Regarding inference thresholds, a score threshold of 0.7 was selected after calibration, suppressing the sensitivity of the pipeline while maintaining high recall. A mask threshold of 0.65 was chosen as it consistently provided a balance between precision and recall while preventing the generation of masks that were excessively loose or tight relative to the ground truth, thereby ensuring accurate diameter measurement. The choice of these thresholds is further discussed in the context of population-level sizing accuracy in Section III-A.

E. Mask Characterization

Following segmentation, each detected mask undergoes a distinct analytical routine to extract an effective agglomerate diameter without modifying the underlying segmentation mask produced by the network. Specifically, the shortest-chord diameter, D_{sc}, is calculated universally for all masks, regardless of whether they are round or elongated. This same algorithm is applied to the manual ground-truth masks, ensuring that all reported diameters are computed identically.

The shortest-chord diameter is computed as follows: the mask centroid is computed from the boundary pixels; the precise contour is extracted using the Marching Squares algorithm [26]; rays are projected from the centroid to the boundary using the Bresenham line algorithm [27] to enumerate valid chords passing through the centroid; the minimum chord length among these candidates is selected as D_{sc}.

Alternative metrics, including area-equivalent (D_A) and perimeter-equivalent (D_P) diameters, were considered but ultimately ruled out. Because area- and perimeter-based metrics integrate the blur-elongated dimension of the masks, they systematically inflate the estimated diameter. Consequently, D_{sc} is adopted as it effectively mitigates this blur-induced overestimation and maintains consistency with the manual baseline.

Figure 4 shows an enlarged close-up comparing the ground-truth manual annotation with the model-generated mask for a representative motion-blurred agglomerate. This visualization explicitly details the extraction of the shortest-chord diameter, demonstrating a qualitative example of the tight alignment between the predicted mask and the manual baseline.

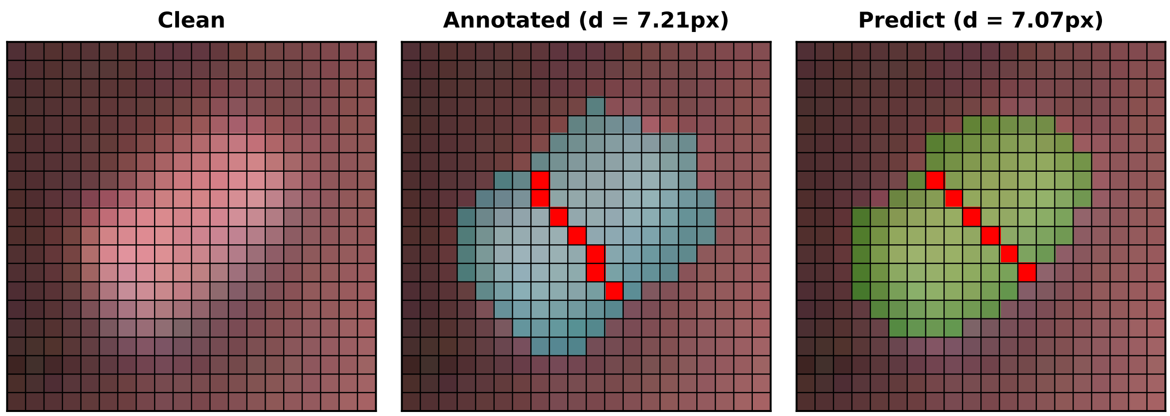


Fig. 4 Enlarged close-up of a representative motion-blurred agglomerate. Left: raw image with no mask overlay. Middle: manual annotation, $D_{sc} = 7.21$ px. Right: model-generated mask, $D_{sc} = 7.07$ px.

III. Results

A. Overall Performance

The primary evaluation of the model focused on its ability to correctly identify agglomerates (detection) and accurately delineate their boundaries (segmentation). Detection performance was quantified using standard metrics: Precision, Recall, and the F1-score. A detected mask was considered a True Positive (TP) if its Intersection over Union (IoU) with a ground truth annotation exceeded 0.5.

A broad qualitative assessment of the model inference capabilities is presented in Fig. 5, highlighting its successful suppression of background noise. The visualization also captures the inherent blur of the raw dataset, stemming from a 15 µs exposure time and mild optical defocus. These artifacts are fully incorporated into the performance evaluation: the shortest-chord algorithm explicitly handles the motion-elongated particles, whereas defocused particles natively contribute to a reduced IoU in their size bin.

To more conveniently illustrate the tight segmentation contours generated around the agglomerates, an extreme close-up is provided in Fig. 4. The application of the confidence threshold (> 0.7) and mask threshold (> 0.65) effectively filters out low-probability artifacts while preserving the geometry of valid particles. A sensitivity check comparing the size distributions obtained from the true-positive detections only against those obtained from all inference results (true positives plus false positives) showed that the two distributions are nearly identical in both count and volume, confirming that the residual false-positive rate does not meaningfully skew the reported distributions. Quantitative results for the model are summarized in Table 2. Because these values evaluate the full population rather than a simple arithmetic mean of the individual pressure columns, they naturally reflect the larger sample size of the 2 MPa subset.

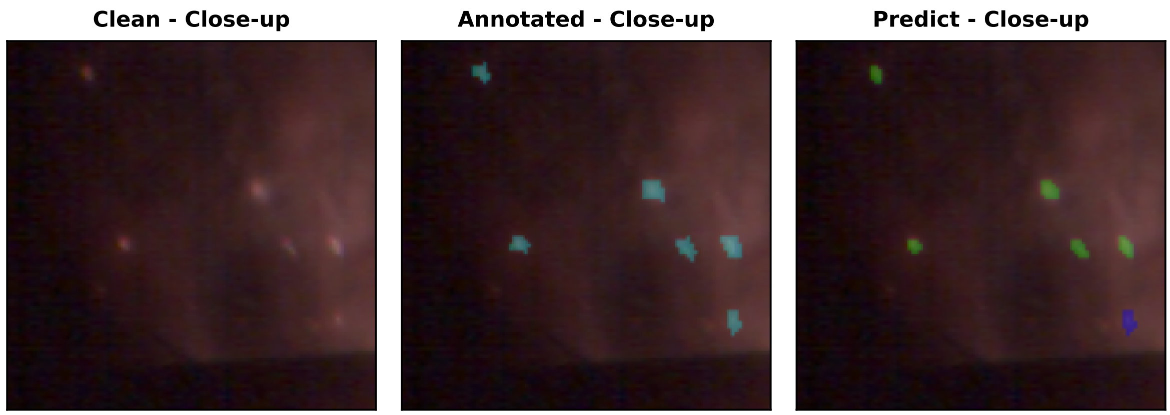


Fig. 5 Segmentation close-up from Fig. 3. Left: raw image. Middle: ground truth annotation. Right: model segmentation output (green denotes true positives and blue denotes false negatives).

The model achieved an F_1-score of 0.82, indicating a strong balance between precision and recall. Per-pressure F_1-score values are 0.82 at 2 MPa, 0.79 at 3 MPa, and 0.82 at 4 MPa, demonstrating consistent detection performance across all examined conditions.

The detection task is explicitly decoupled from the measuring task: the IoU threshold of ≥ 0.5 is used only for binary classification of a detection as a True Positive, while sizing is evaluated independently at the population level. Following standard practice in agglomeration studies, we report the Sauter mean diameter D_{32} and the volume-weighted mean diameter D_{43}, along with their discrepancy ratios (DR) against the manual baseline, defined as:

$$\mathrm{DR}_{32} = \frac{D_{32,\text{inference}}}{D_{32,\text{ground truth}}} \qquad \mathrm{DR}_{43} = \frac{D_{43,\text{inference}}}{D_{43,\text{ground truth}}} \tag{1}$$

A model that perfectly reproduces the manual distributions yields $\mathrm{DR}_{32} = \mathrm{DR}_{43} = 1.0$. Because both metrics are derived from the same shortest-chord algorithm, these ratios strictly isolate the ability of the model to reproduce human-defined boundaries. Moreover, the model demonstrates alignment with the manual baseline, yielding overall DR_{32} of 1.01 and DR_{43} of 0.97. Per-pressure DR_{32} and DR_{43} values are 1.01 and 0.97 at 2 MPa, 1.04 and 1.02 at 3 MPa, and 1.14 and 1.13 at 4 MPa, demonstrating highly consistent sizing at lower pressures with a slight systematic overestimation at the highest evaluated condition.

Table 2 Performance metrics for agglomerate detection and measurement. The "Overall" column reports metrics computed across the pooled dataset.

Metric	Overall	2 MPa	3 MPa	4 MPa
Precision	0.84	0.84	0.84	0.84
Recall	0.80	0.81	0.78	0.80
F1-Score	0.82	0.82	0.79	0.82
DR_{32}	1.01	1.01	1.04	1.14
DR_{43}	0.97	0.97	1.02	1.13

B. Size Distribution

Figure 6 compares the agglomerate diameter distribution inferred by the model against the ground truth. Bin labels indicate bin centers (e.g., the 20 µm tick represents the 10–30 μm bin). Diameters below approximately 20 µm fall at or below the resolvable limit of the imaging system (10–13 μm/pixel) and are subject to significant pixel-quantization uncertainty; a minimum-particle-dimension filter of 4 pixels was applied prior to this analysis to suppress sub-pixel artifacts. The model accurately reconstructs the population statistics, showing a high degree of fidelity in the 40 µm to 200 µm range. A distinct drop-off is observed in the smallest bin (<40 µm), where the model exhibits reduced sensitivity, a known limitation when segmenting small objects represented by a small number of pixels. This deviation, however, is less critical for performance evaluation, since small particles are not as significant as large ones when agglomeration is analyzed, due to their small relative mass fraction, as well as to their reduced effect on the ejected gas flow and their potential to fully combust within the motor. Quantitatively, restricting the ground-truth baseline to particles ≥ 40 µm shifts the median diameter D_{50} only marginally. For instance, the value increases from 111.25 µm to 117.34 µm at 2 MPa (which accounts for the largest fraction of the overall population), with similarly small deviations observed at 3 MPa and 4 MPa. This confirms that the sub-40 µm population has a negligible impact on the total ejected volume.

The volumetric distribution of particles, which is of more interest due to its direct effect on agglomerate mass, is analyzed in Fig. 7. Since volume scales with D^3, small errors in detecting large particles or slight sizing misclassifications can disproportionately affect this metric. Consequently, some bin-to-bin shifting is observed between the model inference and ground truth across the broader 60 µm to 180 µm range, where the model slightly redistributes the volume fraction among adjacent bins. Despite these localized oscillations and a slight under-representation at the extreme upper tail (>200 µm), the model successfully captures the global mass distribution of the agglomerate population.

Image analysis was conducted for different pressures (2 MPa, 3 MPa and 4 MPa), with the volume fraction distribution presented in Fig. 8. For each operating pressure, the median diameter of the model was calculated, and compared to that of the ground truth, as well as to that of the theoretical agglomeration model by Yavor et al. [1]. The inverse relationship between pressure and median volumetric diameter is presented in Table 3 for all three methods, with reasonable agreement between the Mask R-CNN results and the measured data, with the theoretical model [1]

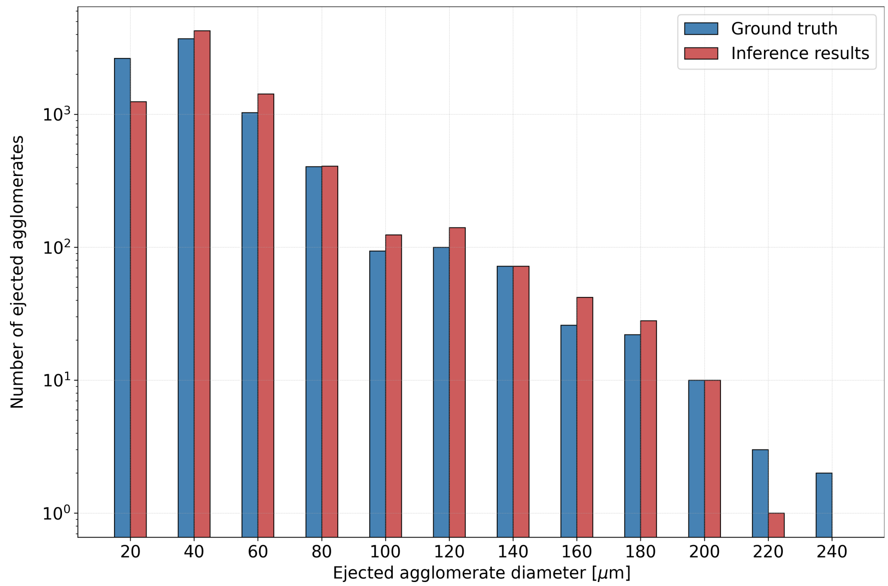


Fig. 6 Number of particles vs. agglomerate diameter: Ground truth vs. segmentation model inference. Note that the y-axis is displayed on a logarithmic scale.

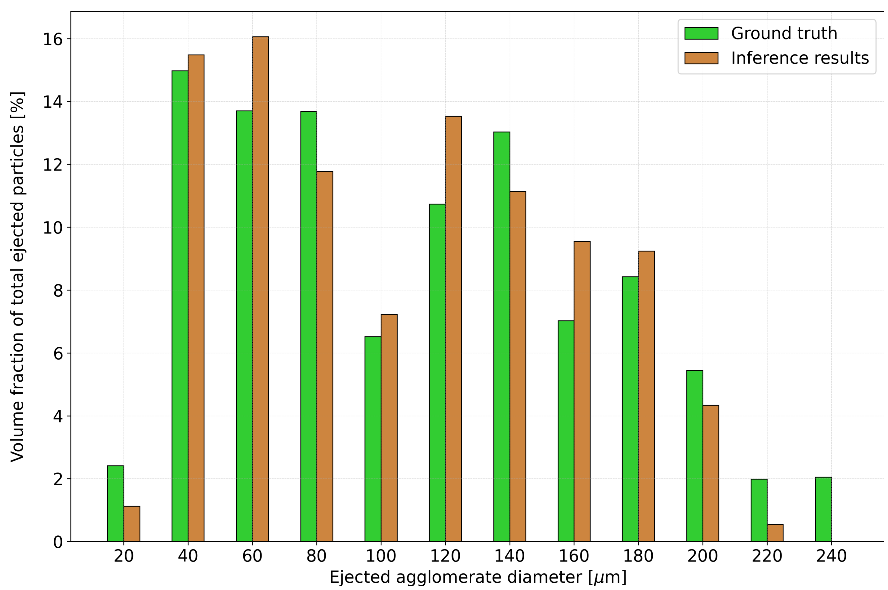


Fig. 7 Size distribution (% of total volume of ejected particles) vs. agglomerate diameter: Ground truth vs. segmentation model inference.

Table 3 Median agglomerate diameter [μm] calculated with various methods for different pressures.

P [MPa]	Mask R-CNN	Manual	Theoretical model [1]
2	112	111	81
3	74	75	78
4	48	48	65

capturing the trend as well.

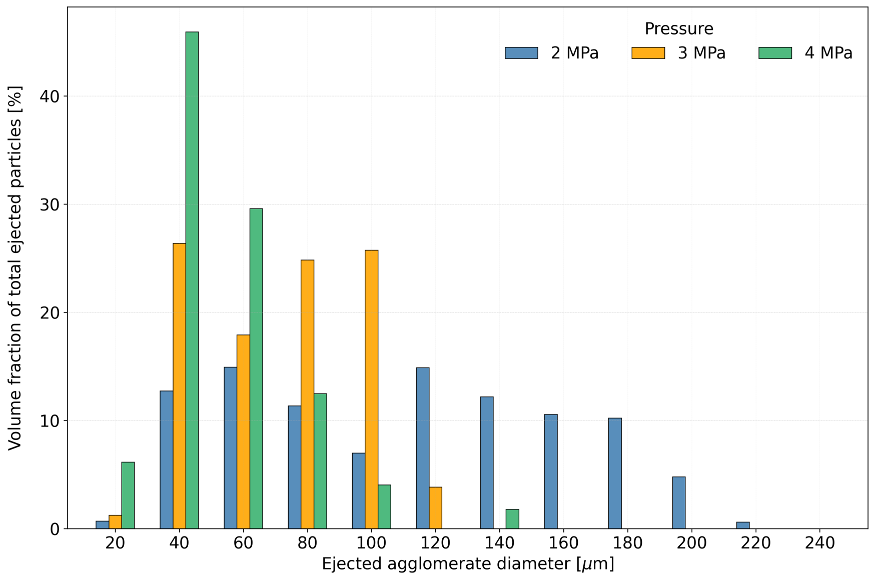


Fig. 8 Size distribution (% of total volume of ejected particles) vs. agglomerate diameter per pressure: Segmentation model inference.

C. Segmentation Performance

The quality of the generated masks was further analyzed by examining the IoU distribution of true positive detections as a function of particle diameter (Fig. 9). There is a clear positive relation between D_{sc} and segmentation quality: the median IoU increases steadily with diameter and stabilizes near 0.80 for the largest bins (160 µm to 200 µm). Performance degrades for particles below 40 µm, where the limited number of pixels available prevents precise boundary delineation; this is an expected geometric behavior, since a single-pixel boundary error disproportionately penalizes the calculated IoU of a small mask compared to a larger one. Finally, because the ground-truth population drops sharply for extreme outliers (only 11 of the 8109 particles exceed 200 µm), the model yields no true positive detections to evaluate in the 220 µm and 240 µm ranges, which is why the statistical analysis naturally truncates at the 200 µm bin.

Shifting from particle-level segmentation to image-level consistency, the robustness of the method is visualized in Fig. 10, which plots the precision and recall for each individual image. 27 images with 10 or fewer annotations were excluded, since the low object count makes the per-image performance metrics inherently unstable. The high density of points in the upper-right quadrant demonstrates that the model performs consistently well on the majority of images.

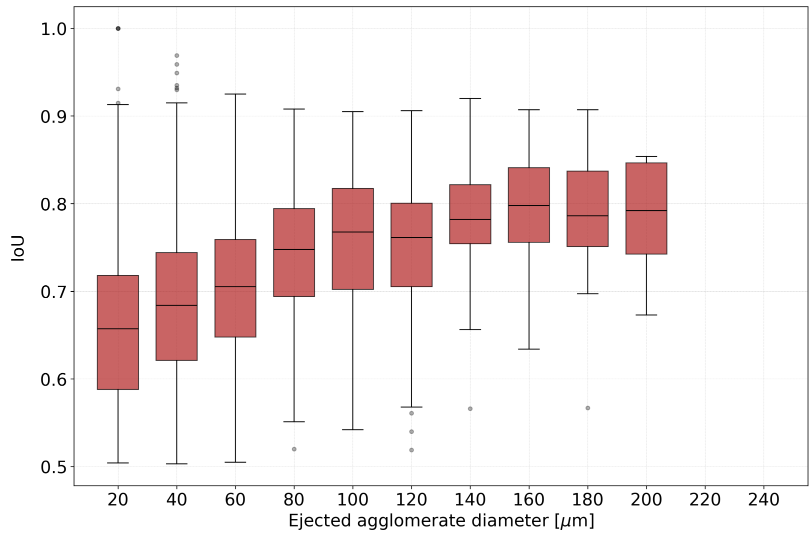


Fig. 9 Distribution of IoU scores across different diameter bins (true positives only), illustrating improved segmentation performance for larger particles.

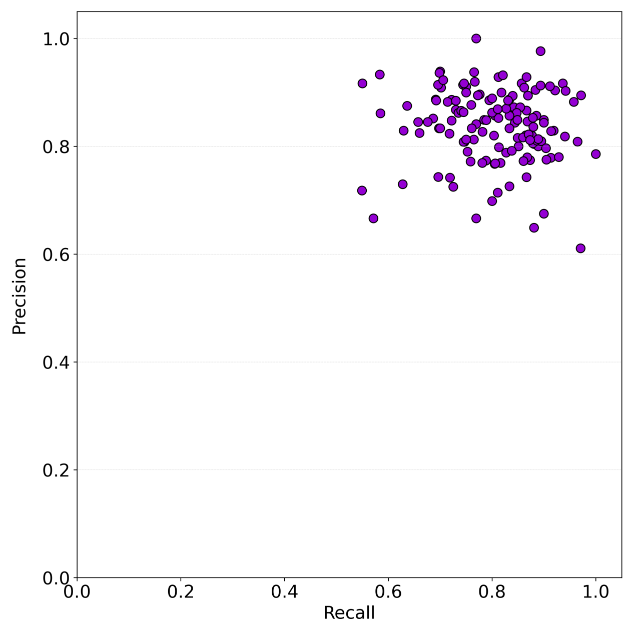


Fig. 10 Per-image precision vs. recall scatter plot. The clustering in the top-right corner indicates consistent high performance across the majority of the dataset.

IV. Discussion

The results demonstrate that the proposed Mask R-CNN pipeline provides a reliable automated alternative to manual annotation for characterizing propellant combustion products. The strong clustering of per-image performance metrics (Fig. 10) indicates that the model generalizes well across the majority of experimental conditions, with performance degradation limited to a small subset of frames exhibiting extreme noise or obscuration.

A. Interpretation of Measurement Accuracy

A critical challenge in automated particle sizing is the dependency of segmentation quality on object resolution. The analysis of IoU versus diameter (Fig. 9) confirms that segmentation fidelity degrades for particles below 40 µm. This is an inherent limitation of the imaging resolution rather than the model architecture; when a particle is represented by only a few pixels, the model performance decreases. However, the impact of these segmentation errors on the total ejected mass estimation is mitigated by the physics of the system. Since volume scales with the cube of the diameter (D^3), the volumetric contribution of these small, poorly segmented particles is less severe.

Furthermore, the reported performance metrics should be interpreted in the context of ground-truth subjectivity. Manual annotation is inherently subject to human error, particularly for small, out-of-focus, or partially occluded particles where annotators may be inconsistent. Manual labeling is also restricted to polygonal approximations, whereas the predicted masks inherently exhibit smoother boundaries that more accurately reflect the true morphology. To strengthen the baseline, the manual annotations were iteratively cross-verified to identify and correct any initially missed or mislabeled particles. The absence of independent physical sizing measurements means that our evaluation primarily reflects agreement with this verified manual baseline rather than absolute metrological accuracy; we therefore report population-level discrepancy ratios, which are substantially more robust to inter-annotator boundary variability than per-particle errors.

B. Comparison with State-of-the-Art

Our methodology addresses specific gaps identified in prior studies. Classical image processing techniques, such as the MSER-based approach and subsequent curvature-based methods [18], rely on handcrafted feature extraction and typically employ simplified equivalent-area fits for diameter estimation. Furthermore, these methods have not been explicitly tested on aluminum particles in solid propellants. In contrast, our proposed deep learning approach leverages learned features from a pre-trained vision model to achieve robust segmentation, validating the resulting physical measurements against a manually annotated dataset.

Quantitatively, Decker et al. reported F_1-scores in the range 0.83–0.90 using shadowgraphy at atmospheric pressure on propellants loaded with inert particles to limit smoke and simplify agglomeration patterns; sizing accuracy was reported in terms of discrete tolerance bins (e.g., 84% of particles within a ±25% tolerance) rather than a single aggregated error metric. Powell et al. [19] compared the inter-method variability of three sizing techniques at atmospheric pressure (reported inter-method variation up to $\approx 30\%$) but did not provide per-method precision or recall metrics against a ground truth, precluding a direct numerical comparison. Despite our shift away from these classical techniques, our pipeline achieves a comparable F_1-score of 0.82 under significantly harsher optical conditions (self-emission imaging at 2–4 MPa with no backlighting), while supplying explicit population-level sizing fidelity via DR_{32} and DR_{43}. We note that a true apples-to-apples comparison across methods would require applying each to the identical dataset, an important direction for future work.

In the context of recent deep learning advancements, our approach strikes a balance between accessibility and precision. While studies utilizing high-end microscopic imaging [22, 23] achieve higher raw segmentation scores, they require instrumentation that is often impractical for standard firing tests. Conversely, real-time tracking models, like YOLO11 [24] prioritize inference speed. While valuable for tracking, real-time architectures often approximate boundaries (bounding boxes or coarse masks) that are insufficient for precise morphological measurements. By utilizing a heavier ResNet-50 backbone and introducing the post-processing shortest-chord metric, our method prioritizes physical accuracy over real-time processing, making it more suitable for detailed post-experiment material characterization.

C. Methodological Contributions

The adoption of the shortest-chord diameter, D_{sc}, provides a practical refinement tailored for sizing motion-blurred particles, a common challenge in high-speed combustion videography. By naturally resisting the blur-elongated dimensions of the masks, D_{sc} actively mitigates systematic diameter inflation. Applying this single, robust metric universally across all detections ensures a continuous and stable evaluation. Furthermore, because this identical sizing routine is applied to the manual ground-truth masks, the resulting population-level evaluation via DR_{32} and DR_{43} effectively isolates and measures the consistency of the model against the human baseline, rather than claiming absolute metrological accuracy.

V. Conclusion

This work shows that a Mask R-CNN-based segmentation pipeline can serve as a practical and reliable alternative to manual annotation for the characterization of propellant combustion products. Its performance remains stable across most experimental conditions, indicating that robust automated analysis is achievable even in the optically challenging environment of high-speed combustion imaging. Beyond reducing the labor and subjectivity associated with manual labeling, the approach enables a more consistent basis for extracting physically meaningful particle statistics from large image sets.

A key outcome is that the main limitation of the method lies not in the segmentation architecture itself, but in the finite resolution of the imaging system. Accuracy decreases for small particles, yet the influence of these errors on bulk mass-related estimates is low due to their limited volumetric contribution. Moreover, one must note that manual annotation is not an absolute standard, but a human-interpreted baseline with its own uncertainty. Framing performance in terms of population-level discrepancy therefore provides a more meaningful assessment of measurement fidelity than strict per-particle agreement alone.

The study also demonstrates the value of the shortest-chord diameter as a sizing metric for motion-blurred particles. By reducing sensitivity to blur-driven elongation, this metric mitigates systematic diameter inflation and ensures consistent morphological evaluation. Taken together, these results establish a useful framework for post-experiment particle analysis that balances accessibility, robustness, and measurement relevance. Future progress should focus on validation against independent physical sizing measurements and direct benchmarking against alternative methods on identical datasets, which would further define the metrological limits and practical advantages of the approach.

References

[1] Yavor, Y., Gany, A., and Beckstead, M. W., "Modeling of the Agglomeration Phenomena in Combustion of Aluminized Composite Solid Propellant," *Propellants, Explosives, Pyrotechnics*, Vol. 39, No. 1, 2014, pp. 108–116. https://doi.org/10.1002/prep.201300073.

[2] Crump, J. E., Prentice, J. L., and Kraeutle, K. J., "Role of the Scanning Electron Microscope in the Study of Solid Propellant Combustion: II. Behavior of Metal Additives," *Combustion Science and Technology*, Vol. 1, No. 3, 1969, pp. 205–223. https://doi.org/10.1080/00102206908952201.

[3] Gany, A., Caveny, L. H., and Summerfield, M., "Aluminized Solid Propellants Burning in a Rocket Motor Flowfield," *AIAA Journal*, Vol. 16, No. 7, 1978, pp. 736–739. https://doi.org/10.2514/3.60959.

[4] Glotov, O. G., Yagodnikov, D. A., Vorob'ev, V. S., Zarko, V. E., and Simonenko, V. N., "Ignition, Combustion, and Agglomeration of Encapsulated Aluminum Particles in a Composite Solid Propellant. II. Experimental Studies of Agglomeration," *Combustion, Explosion, and Shock Waves*, Vol. 43, No. 3, 2007, pp. 320–333. https://doi.org/10.1007/s10573-007-0045-y.

[5] DeLuca, L. T., Galfetti, L., Colombo, G., Maggi, F., Bandera, A., Babuk, V. A., and Sinditskii, V. P., "Microstructure Effects in Aluminized Solid Rocket Propellants," *Journal of Propulsion and Power*, Vol. 26, No. 4, 2010, pp. 724–732. https://doi.org/10.2514/1.45262.

[6] Beckstead, M. W., "An Overview of Aluminum Agglomeration Modeling," *50th Israel Annual Conference on Aerospace Sciences*, 2010.

[7] Crump, J. E., "Aluminum Particle Combustion in Composite Solid Propellants," *18th Joint Propulsion Conference*, Chemical Propulsion Information Agency, Silver Spring, 1966, pp. 321–329. https://doi.org/10.2514/6.1982-1110.

[8] Cohen, N. S., "A Pocket Model for Aluminum Agglomeration in Composite Propellants," *AIAA Journal*, Vol. 21, No. 5, 1983, pp. 720–725. https://doi.org/10.2514/3.8139.

[9] Gallier, S., "A Stochastic Pocket Model for Aluminum Agglomeration in Solid Propellants," *Propellants, Explosives, Pyrotechnics*, Vol. 34, No. 2, 2009, pp. 97–105. https://doi.org/10.1002/prep.200700260.

[10] Gany, A., and Caveny, L. H., "Agglomeration and Ignition Mechanism of Aluminum Particles in Solid Propellants," *Symposium (International) on Combustion*, 1978, pp. 1453–1461. https://doi.org/10.1016/S0082-0784(79)80137-X.

[11] Yavor, Y., Rosenband, V., and Gany, A., “Reduced Agglomeration in Solid Propellants Containing Porous Aluminum,” *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, Vol. 228, No. 10, 2014, pp. 1857–1862. https://doi.org/10.1177/0954410013495638.

[12] Yavor, Y., Rosenband, V., and Gany, A., “Reduced Agglomeration Resulting from Nickel Coating of Aluminum Particles in Solid Propellants,” *International Journal of Energetic Materials and Chemical Propulsion*, Vol. 9, No. 6, 2010. https://doi.org/10.1615/IntJEnergeticMaterialsChemProp.2011001385.

[13] Liu, T. K., “Experimental and Model Study of Agglomeration of Burning Aluminized Propellants,” *Journal of Propulsion and Power*, Vol. 21, No. 5, 2005, pp. 797–806. https://doi.org/10.2514/1.11988.

[14] Cohen, O., Michaels, D., and Yavor, Y., “Agglomeration in Composite Propellants Containing Different Nano-Aluminum Powders,” *Propellants, Explosives, Pyrotechnics*, Vol. 47, No. 9, 2022, p. e202100320. https://doi.org/10.1016/j.dt.2022.05.016.

[15] Sharabi, N., Abdalla, Y., and Yavor, Y., “Effect of Coated Nano-Aluminum Powders on the Agglomeration in Solid Propellant,” *9th European Conference for Aeronautics and Space Sciences (EUCASS)*, Lille, France, 2022. https://doi.org/10.13009/EUCASS2022-7246.

[16] Tu, C. Y., Chen, X., Li, Y. K., Zhang, B. C., and Zhou, C. S., “Experimental Study of Al Agglomeration on Solid Propellant Burning Surface and Condensed Combustion Products,” *Defence Technology*, Vol. 26, 2023, pp. 111–122. https://doi.org/10.1016/j.dt.2022.05.016.

[17] Xiao, L., Pang, W., Qin, Z., Li, J., Fu, X., and Fan, X., “Cluster Analysis of Al Agglomeration in Solid Propellant Combustion,” *Combustion and Flame*, Vol. 203, 2019, pp. 386–396. https://doi.org/10.1016/j.combustflame.2018.12.032.

[18] Decker, T. G., Devillers, R. W., and Gallier, S., “Detecting Agglomeration Patterns on Solid Propellant Surface via a New Curvature-Based Multiscale Method,” *Acta Astronautica*, Vol. 206, 2023, pp. 123–132. https://doi.org/10.1016/j.actaastro.2023.02.020.

[19] Powell, M. S., Gunduz, I. W., Shang, W., Chen, J., Son, S. F., Chen, Y., and Guildenbecher, D. R., “Agglomerate Sizing in Aluminized Propellants Using Digital Inline Holography and Traditional Diagnostics,” *Journal of Propulsion and Power*, Vol. 34, No. 4, 2018, pp. 1002–1014. https://doi.org/10.2514/1.B36859.

[20] Nugue, M., Roche, J.-M., Le Besnerais, G., Trottier, C., Devillers, R. W., Pichillou, J., Chan-Hon-Tong, A., Boulch, A., and Hurmane, A., “Scaling Up Information Extraction from Scientific Data with Deep Learning,” *AerospaceLab Journal*, 2020, pp. AL15–07. https://doi.org/10.12762/2020.AL15-07.

[21] Sportillo, A., Airiau, M., Devillers, R. W., and Chan-Hon-Tong, A., “Improving Alumina-Cap Characterisation in Aluminized Solid-Rocket Motors with Computer Vision,” *IAF Space Propulsion Symposium*, 2024, pp. 301–315. https://doi.org/10.52202/078371-0035.

[22] Frei, M., and Kruis, F. E., "Image-Based Size Analysis of Agglomerated and Partially Sintered Particles via Convolutional Neural Networks," *Powder Technology*, Vol. 360, 2020, pp. 324–336. https://doi.org/10.1016/j.powtec.2019.10.020.

[23] Wang, Y., Zhang, H., Zhuo, Z., Shen, B., Wu, S., Ao, W., Chen, D., Wu, Y., and Wu, X., "Aluminum Droplet, Oxide Cap and Flame Segmentation in Burning Al/AP Propellants by Combining YOLOv7 and Two-Stage Cluster," *Measurement*, Vol. 227, 2024, p. 114264. https://doi.org/10.1016/j.measurement.2024.114264.

[24] Xue, X., Gao, H., Sun, Z., and Liu, J., "Segmentation and Tracking: A Deep Learning-Based Method for Analyzing Dynamic Features of Aluminum Agglomerates," *Powder Technology*, Vol. 455, 2025, p. 120785. https://doi.org/10.1016/j.powtec.2025.120785.

[25] Brooks, J., "COCO Annotator: A web-based image annotation tool," https://github.com/jsbroks/coco-annotator, 2019.

[26] Lorensen, W. E., and Cline, H. E., "Marching Cubes: A High Resolution 3D Surface Construction Algorithm," *ACM SIGGRAPH Computer Graphics*, Vol. 21, No. 4, 1987, pp. 163–169. https://doi.org/10.1145/37402.37422.

[27] Bresenham, J., "A Linear Algorithm for Incremental Digital Display of Circular Arcs," *Communications of the ACM (CACM)*, Vol. 20, No. 2, 1977, pp. 100–106. https://doi.org/10.1145/359423.359432.