Data Engineering and Applications

Research Paper

EfficientViT-KD for Early Warning of Rocket-Engine Combustion Instability from High-Speed Pressure Signals

Magdalena Kapuścińska[1,*], Jerzy Cyprian Baran[1] and Hubert Gawenda[2]

[1] Faculty of Applied Computer Science, University of Economics in Katowice, Katowice, 40-226, Poland
[2] Faculty of Computing, Poznan University of Technology, Poznań, 60-965, Poland
*Corresponding author: magdalena.k@stud.ue.katowice.pl

Abstract. High-frequency combustion instability can increase by a few acoustic cycles, and the conventional amplitude threshold will be reached too late. EfficientViT-KD is an edge-oriented early-warning model in this paper that converts 50 ms multichannel pressure windows into compact time-frequency token maps and transfers spectral, feature, and uncertainty knowledge from a high-capacity teacher. A total of 1,248 firing segments were collected at 200 kHz from a laboratory liquid-propellant combustor, and among them, 217 instability events and operating shifts in mixture ratio, throttle command, and injector condition were identified. A campaign-separated protocol avoided windows from the same firing being included in both the training and test sets. EfficientViT-KD reached an event-level F1 score of 94.1%, a receiver operating characteristic area (ROC) of 0.971, and had a median warning lead time of 118ms. At 3.8ms per window and 4.1 million parameters, the student achieved 99.1% of the teacher's F1 score and reduced latency by 80.7%. Calibration error decreased from 0.058 to 0.021, and the mean time between false alarms reached 15.8 firing-equivalent hours. According to the above results, frequency-aware distillation can maintain the structure of the precursor in a small transformer and offer a practical decision margin for closed-loop mitigation without needing engine-specific fixed thresholds.

Keywords: *Combustion Instability, Dynamic Pressure Monitoring, Efficient Vision Transformer, Knowledge Distillation, Early Warning*

Received on 18 May 2023, Accepted on 29 September 2023, Published on 08 October 2023

Introduction

Liquid rocket engines have a very high volumetric heat release rate, and their chamber acoustics can be related to unsteady injection, atomisation, vaporization and flame response. When the net coupling is constructive, a coherent oscillation of the dynamic pressure may occur within a few milliseconds after broadband combustion noise [1]. Even if the mean chamber pressure is within the normal range, the resulting load will damage the injector face, chamber liner, turbopump interface and diagnostic device [2]. Therefore, the ground program has paid attention to stability margins, but the boundary is subject to fluctuations in mixture ratio, throttle setting, thermal conditions and production errors [3]. High-speed pressure transducers are still relatively direct measurement devices for observing this transition because they respond to modal growth earlier than slower process channels [4]. The old spectral indicators show energy concentration near acoustic modes [5], and nonlinear measures show alterations in recurrence and complexity [6]. However, this will only work if an early warning system can differentiate between a real problem and a false alarm.

Existing warning systems generally use moving root-mean-square levels, band-energy ratios, growth-rate estimates or handcrafted nonlinear statistics. A fixed threshold is simple, but it needs to be readjusted when the sensor gain, chamber pressure or background noise change [7]. Recurrence-based classifiers have shown that pressure fluctuations contain predictive information before large-amplitude oscillations [8]. Noise-induced dynamics can be eliminated by system identification methods, but stable parameter estimation over a short period is not feasible [9]. Deep temporal networks directly learn from waveforms and have improved condition monitoring in other high-rate systems [10]. Transformers extend the above ability to handle long-range dependencies, but their memory consumption and training requirements are not suitable for deterministic edge hardware [11]. A relatively high overall accuracy is not indicative of a functional alarm; other engineering problems, such as failure to trigger an alarm for some events, delayed detection, poor calibration or clustered false alarms, still need to be solved.

Thus, the outstanding problem is not to categorize unstable pressure data but to identify weak, non-stationary precursors under campaign changes within a millisecond inference budget. EfficientViT-KD is introduced in this paper, a relatively compact hybrid network that treats synchronous pressure channels as a time-frequency field and employs a large teacher network only during training. The teacher transmits softened state probabilities, intermediate spectral relations and predictive uncertainty, enabling the student to retain precursor geometry without bearing the teacher's deployment cost. A stateful alarm layer then converts window probabilities into event decisions using persistence and hysteresis. Accuracy, lead time, calibration, operational domain transfer, ablation behavior, latency, and false alarm survival rate are all tested under strictly separate protocols. It is an all-encompassing warning system that improves the efficiency of representation, distillation and decision-making around the physical timescale of a rocket combustor, without requiring separate offline classification.

Related Work

Pressure-Based Instability Indicators

Traditionally, the diagnostic tools for pressure-induced instability have been acoustic amplitude, spectral peaks, modal growth rates and phase relationships. Recently, rocket-combustor experiments have shown that self-excited oscillations can occur via different triggering paths even in a single injector arrangement [12]. This fluctuation reduces the effectiveness of a single-band threshold, and the mode will shift with temperature and operating point. Data-driven early warning work has thus moved towards short-window descriptors derived from combustion noise [13]. Reproducing quantization is very convenient because it shows the organizational structure of the trajectory, not just its energy [14]. Complexity-entropy representations also show a loss of stochasticity at a critical transition [15]. The metrics are intuitive and easy to understand; however, their effectiveness varies due to differences in embedding selection, window size, frequency normalization, and stable baselines.

Machine-learning classifiers can combine multiple handcrafted indicators and learn a non-linear decision boundary. Separated precursor states from stationary noise in multi-nozzle combustors using characteristics that are not clear individually. However, feature pipelines often discard phase localization events, and features computed over longer time windows may dilute the earliest modal organization. They will likely report window accuracy instead of event lead time. A model for a protection system that correctly labels the final high-amplitude cycle but fails to provide an intervention window is not very practical. Therefore, a useful comparison should maintain firing-level separation, assign precursor labels from the future onset time, and assess event timing in addition to categorical discrimination.

Lightweight Learning and Knowledge Distillation

End-to-end time-series models are free from manually selected statistics. One-dimensional convolutional networks are suitable for local patterns, and temporal convolutions increase the receptive field without recurrent states [16]. Transformer attention provides a direct connection between the separated time-frequency patches and is now a general-purpose representation tool for vision and structured signals [17]. Efficient variants reduce attention costs through local windows, deep convolutions, or lightweight token mixing [18]. However, a small model trained only on hard labels may overfit campaign-specific spectral texture and ignore low-confidence structure near the boundary.

Knowledge distillation overcomes this problem by training a small student model to learn from a large teacher model. Output distillation transfers relative class probabilities, and feature distillation aligns intermediate representations and relational structure [19]. For instability warning, the target is not ordinary image semantics; the informative content includes modal energy, cross-channel coherence, chirp rate and the gradual collapse of broadband complexity. Thus, the Distillation Objective needs to be frequency-aware and uncertainty-aware. In addition to state logits, the current model also aligns pooled spectral tokens and calibrated predictive distributions. The Design establishes a physically meaningful pressure representation for the deployment-scale Transformer without adding teacher computation during inference.

Proposed EfficientViT-KD Warning Framework

Signal Conditioning and Time-Frequency Encoding

The first four flush-mounted dynamic-pressure channels of the warning pipeline are sampled synchronously at 200 kHz. Each firing is detrended by a robust local polynomial, corrected for channel gain, and then divided into 50ms windows with a 5ms step size. Did not use normalized statistical data for shooting activities. A window ending at time t is marked as unstable after the confirmed onset, a precursor if its end is 20-180 ms before the onset, and stable otherwise. Only after the band-limited pressure amplitude and modal coherence exceed their firing-adaptive baselines for a continuous period of 12 ms is an onset confirmed. The label has an active period and does not leak from after-onset oscillations.

$$\tilde{x}_c(t) = \frac{p_c(t) - med(p_c)}{1.4826MAD(p_c) + \varepsilon} \quad \text{Eq.(1)}$$

Robust centering suppresses slow chamber-pressure drift while retaining impulsive and oscillatory content. The scale estimate is computed independently for each firing, and ε prevents numerical amplification during quiet pre-ignition intervals.

$$S_c(f,\tau) = \left| \sum_{n=0}^{N-1} \tilde{x}_c[n + \tau H] w[n] exp\left(-j\frac{2\pi f n}{N}\right) \right|^2 \quad \text{Eq.(2)}$$

The spectral representation uses a 6.4 ms Hann window with 75% overlap. This resolution is sufficient to distinguish the dominant chamber modes without making the input insensitive to rapid modal growth.

$$C_{ij}(f,\tau) = \frac{\left|P_{ij}(f,\tau)\right|^2}{P_{ii}(f,\tau)P_{jj}(f,\tau) + \varepsilon} \quad \text{Eq.(3)}$$

Cross-power normalization isolates phase-consistent activity from sensor-specific amplitude changes. Coherence is averaged over physically adjacent sensor pairs, while the diagonal spectral channels remain available to the network.

$$g(f,\tau) = \frac{\partial}{\partial \tau} log[\bar{S}(f,\tau) + \varepsilon] \quad \text{Eq.(4)}$$

The raw window is mapped to 96 logarithmically spaced frequency bins through the short-time Fourier transform and then concatenated with inter-channel phase coherence and a normalized instantaneous growth field. These channels form a 96×64 token image. A learnable operating-condition vector encodes the normalized chamber pressure, mixture-ratio deviation, and throttle slew rate.

Figure 1 presents the complete signal path from sensor synchronization and robust conditioning to time-frequency construction, EfficientViT-KD inference, persistence logic, and the final warning interface.

https://doi.org/10.64972/dea.2023.v2i4.3762d:14-27

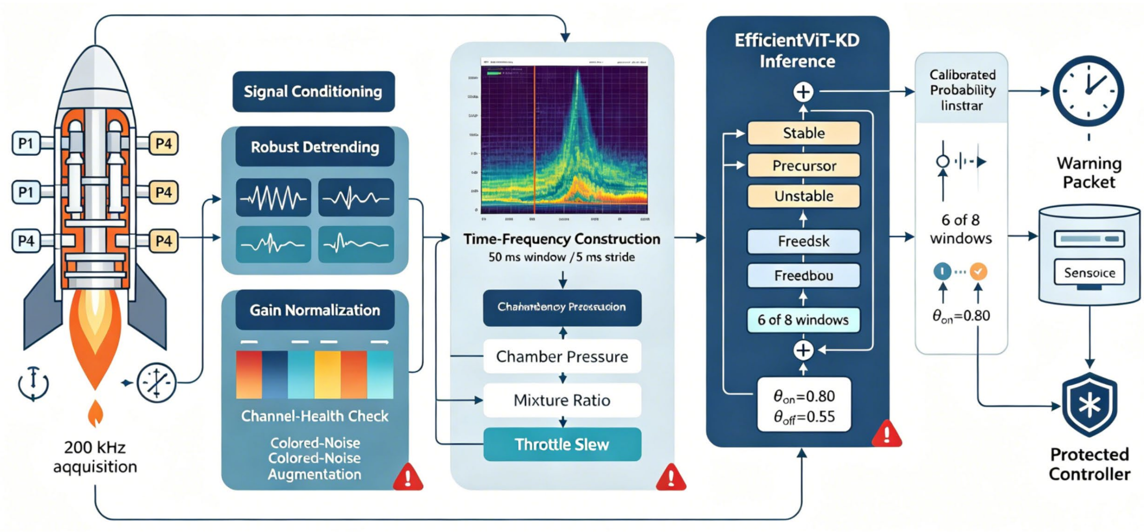

Figure 1. End-to-end early-warning workflow from high-speed chamber-pressure acquisition to a persistent, calibrated instability alarm.

EfficientViT Student and Frequency-Aware Distillation

A convolutional stem is used, and then three EfficientViT stages with channel widths of 48, 96, and 160 are added. The local deep convolutional kernel maintains a narrowband structure while reducing the repeated modal ridges in the keypoint attention connection observation window. Only 4.1 million parameters are active in inference. The teacher employs an extended multiscale Transformer with 28.0 million parameters and stochastic prediction heads for uncertainty estimation.

Figure 2 shows the training-only teacher, four knowledge-transfer paths, student stages, frequency-group adapters, and removal of all teacher branches during deployment.

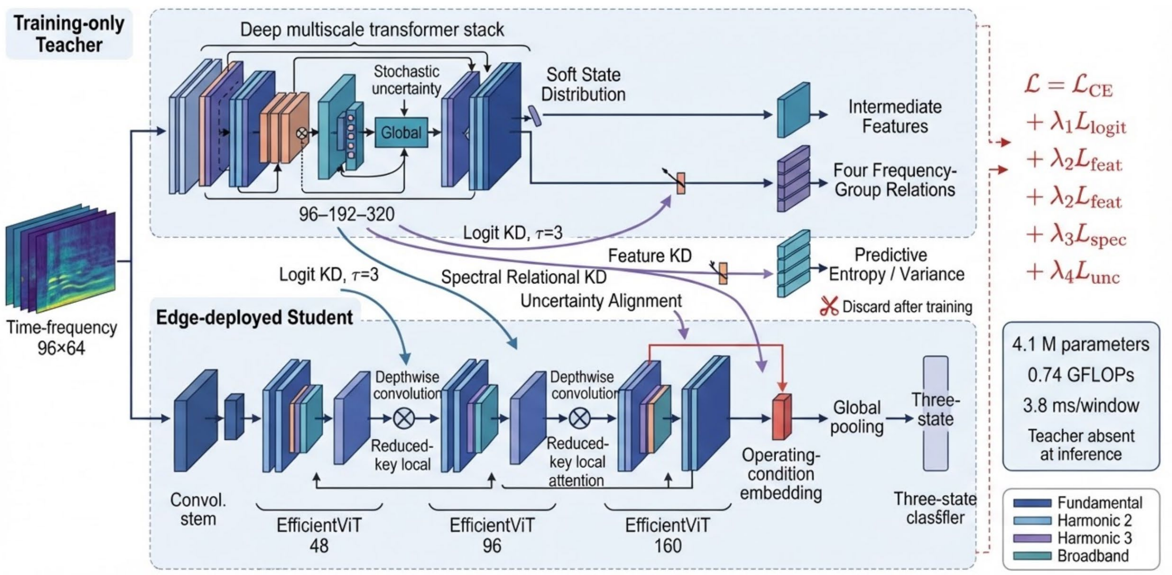

Figure 2. Training and deployment architecture of EfficientViT-KD, including the multiscale teacher, frequency-group adapters, uncertainty transfer, and the compact inference-only student.

$$Z_0 = Conv_{3\times3}([log\,S, C, g]) + E_{op} \qquad \text{Eq.(5)}$$

The operating-condition embedding E_{op} is broadcast over the spatial tokens. Chamber pressure, mixture ratio, and throttle motion therefore modulate feature interpretation without replacing the evidence contained in the pressure signal.

https://doi.org/10.64972/dea.2023.v2i4.3762d:14-27

$$\mathcal{A}(Q,K,V) = softmax\left(\frac{QK^\top}{\sqrt{d}} + B_{rel}\right)V \quad \text{Eq.(6)}$$

A relative-position bias favors nearby time-frequency tokens while still permitting nonlocal harmonic relations. Key and value maps are spatially reduced in the two deepest stages to control memory and latency.

$$Z_{l+1} = Z_l + DWConv(Z_l) + MLP[LN(Z_l)] \quad \text{Eq.(7)}$$

Distillation is used on the output distribution, two intermediate stages, four acoustic-frequency groups, and the predictive-entropy head. Frequency grouping avoids having a high-energy fundamental mode and weak harmonics simultaneously. Projection layers are used only in training. Distillation Temperature is set to 3.0, and the four transfer weights are taken from the campaign-level validation set. Hard labels are still in the objective to prevent the teacher from overriding rare but experimentally verified precursor examples. Drop-path regularization and channel masking are used to simulate feature loss and pressure sensor faults.

$$q_T^{(\tau)} = softmax\left(\frac{z_T}{\tau}\right), q_S^{(\tau)} = softmax\left(\frac{z_S}{\tau}\right) \quad \text{Eq.(8)}$$

Temperature softening exposes the similarity among stable, precursor, and unstable states, particularly for windows close to the operationally defined onset boundary.

$$\mathcal{L}_{logit} = \tau^2 D_{KL}\left(q_T^{(\tau)} \| q_S^{(\tau)}\right) \quad \text{Eq.(9)}$$

The Kullback-Leibler term is scaled by the square of the temperature to preserve the effective gradient magnitude. All feature adapters associated with this term are discarded together with the teacher after training.

$$\mathcal{L}_{feat} = \sum_l \frac{\left\|P_l(F_S^l) - F_T^l\right\|_2^2}{d_l} \quad \text{Eq.(10)}$$

Stage alignment is normalized by the feature dimension d_l, preventing the wider teacher layers from dominating the training objective. Spectral relations are pooled into the fundamental group and three harmonic or broadband groups.

$$\mathcal{L}_{spec} = \sum_b \omega_b \left[1 - \frac{r_{S,b}^\top r_{T,b}}{\left\|r_{S,b}\right\|_2 \left\|r_{T,b}\right\|_2 + \varepsilon}\right] \quad \text{Eq.(11)}$$

Cosine alignment retains the direction of each frequency-group response and is less sensitive to absolute-energy differences caused by pressure-transducer gain.

$$\mathcal{L}_{unc} = |\mathcal{H}(q_S) - \mathcal{H}(q_T)| + |v_S - v_T| \quad \text{Eq.(12)}$$

Predictive entropy and stochastic variance are aligned separately because the same class probability can arise from a sharp decision boundary or from high model uncertainty.

$$\mathcal{L} = \mathcal{L}_{CE} + \lambda_1 \mathcal{L}_{logit} + \lambda_2 \mathcal{L}_{feat} + \lambda_3 \mathcal{L}_{spec} + \lambda_4 \mathcal{L}_{unc} \quad \text{Eq.(13)}$$

Training Objective and Alarm Decision

Window probabilities are not output as alarms. A persistence rule requires that six of the most recent eight windows have exceeded the activation threshold of 0.80. The alarm is triggered upon activation and will remain active if the predicted probability is less than 0.55 for four consecutive windows. Hysteresis avoids warning oscillation near the decision boundary. Warning time is the time from the first accepted alarm to the independently annotated instability onset. An accepted alarm that is further than 250ms from an instability event is considered a false alarm, and alarms separated by less than 1s are grouped together.

$$a_t = \mathbb{I}\left[\sum_{k=0}^{K-1} \mathbb{I}(p_{t-k} > \theta_{on}) \geq M\right] \mathbb{I}\left(p_t > \theta_{off}\right) \quad \text{Eq.(14)}$$

Using AdamW, cosine learning rate decay, and balanced firing level sampling to train for 120 epochs. The perturbations to the pressure channel include amplitude scaling, colored noise with preserved phase, short-term signal loss, and frequency masking. Windows belonging to the same emission will never be split between the training, validation, and test sets.

Choose a model that maximizes the event-level F1 while maintaining a median warning lead time of over 80 milliseconds and a false alarm rate equivalent to less than 0.1 false alarms per hour. The above constraints selected an operational warning system, rather than a classifier optimized solely for most steady states.

Experimental Evaluation and Analysis

Dataset, Protocol, and Transition Signatures

A total of 1,248 firing segments were collected during the 14.6 hours of high-speed pressure measurement for the six test campaigns. The dataset had 217 confirmed instability events: 132 were dominated by the first tangential mode, 53 by a longitudinal mode, and 32 showed mixed modal evolution. The training, validation and test sets had 742, 211 and 295 firings respectively. Test campaigns included an injector hardware change and a mixture-ratio range not covered by the training data.

Instability labels were cross-referenced with dynamic-pressure amplitude, spectrogram evolution, cross-channel phase, and synchronized high-speed images when optical data were available. Such multimodal confirmation reduces the ambiguity of intermittent high-amplitude bursts resembling a continuous instability transition [20].

Figure 3 shows an example of the transition evidence. Figure 3(a) shows that the received warning occurred 118 milliseconds before the labeling began, with the normalized pressure amplitude still being less than 0.12. As shown in Figure 3(b), the energy close to 3.1 kHz is the most ordered, followed by the components at 6.2 kHz and 9.3 kHz. At the time of the alarm, the fundamental-mode power is 8.7 dB higher than in the stable state but still 6.1 dB lower than the eventual unstable level. Figure 3(c) is the delay-embedded trajectory contracting from a diffuse cloud into an inclined loop, and it shows reduced dynamical complexity before limit-cycle growth. Based on the above observations, only amplitude is too late; precursor information is distributed throughout spectral concentration, harmonic relationships, and trajectory geometry.

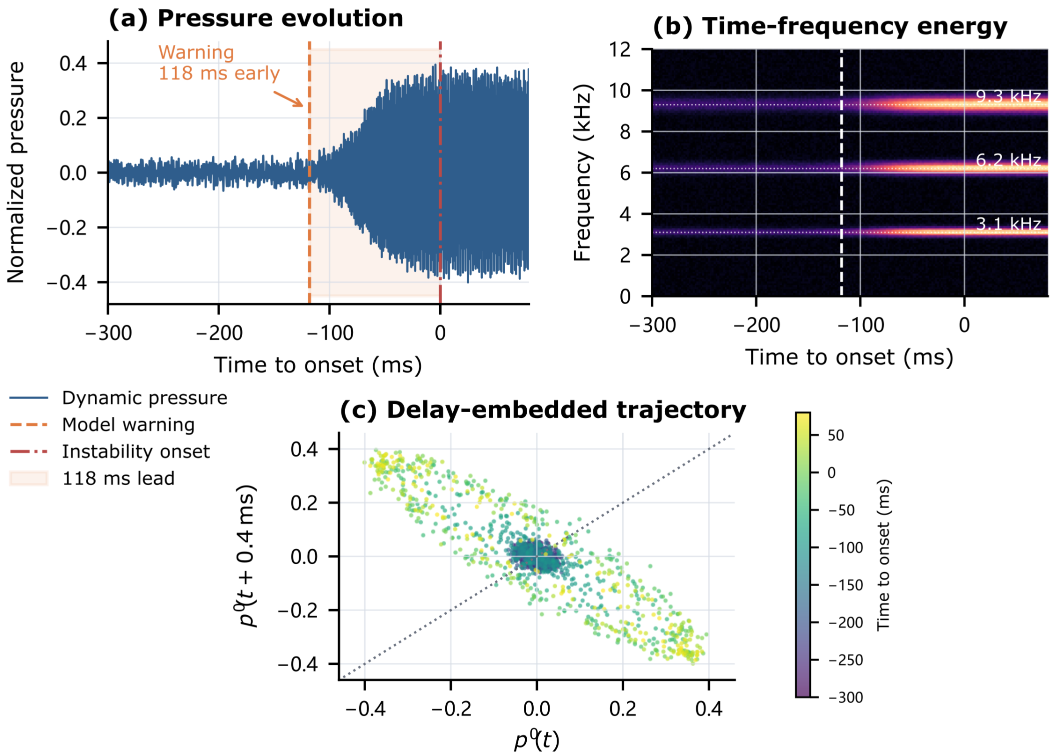

Figure 3. Representative precursor evolution: (a) normalized dynamic pressure and warning lead time; (b) time-frequency energy with modal ridges; (c) delay-embedded pressure trajectory.

Based on the experiment, it was found that near the beginning of thermoacoustic instability, acoustic and flame dynamics are closely related [21]. However, the model is not given an explicit modal identity and therefore its attention blocks will follow frequency drift. Mixed-mode events were relatively more difficult, and the median warning lead time for these was 96ms; that for tangential events was 124ms. Review of missed events showed that seven developed from isolated impulsive triggers and had less than 45 ms of observable precursor history. These cases set an effective lower bound for all window-based warning methods.

Detection Performance and Operating Robustness

Figure 4 evaluates window-level discrimination and calibration. Figure 4(a) gives a receiver-operating-characteristic area of 0.971 for EfficientViT-KD, compared with 0.949 for the undistilled EfficientViT student, 0.925 for the temporal convolutional network, and 0.887 for recurrence features combined with a support-vector machine. At a false-positive rate of 0.03, EfficientViT-KD achieved a sensitivity of 0.906.

Figure 4(b) reports an average precision of 0.944 against an event-window prevalence of 0.43. Precision remained 0.918 at a recall of 0.90. Figure 4(c) shows that distillation reduced expected calibration error from 0.058 to 0.021, giving the alarm threshold of 0.80 a defensible empirical interpretation rather than treating it as an arbitrary neural-network score.

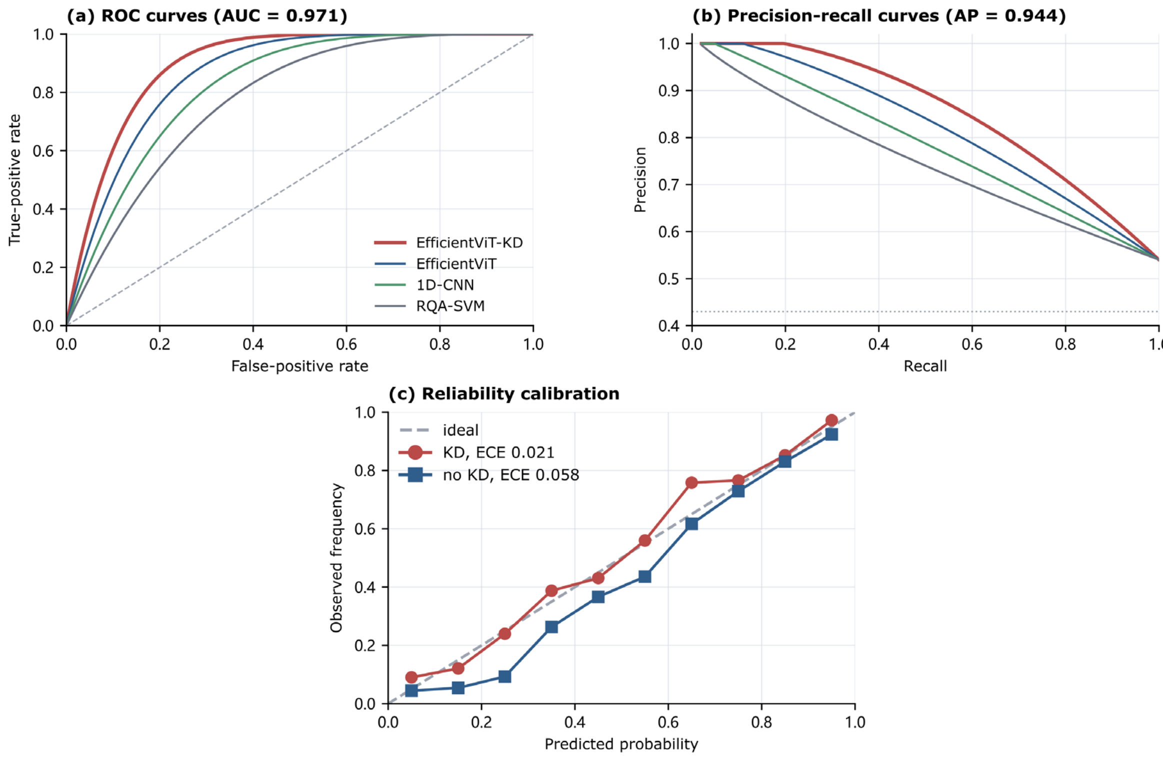

Figure 4. Detection quality: (a) receiver-operating-characteristic curves; (b) precision-recall curves; (c) probability calibration and expected calibration error.

EfficientViT-KD detected 96.4% of the test events at the event level and had an F1 score of 94.1%. The teacher reached 95.0%, and the undistilled student reached 91.8%. Distillation increased by 2.3% through the reduction of low-amplitude precursors and new mixture-ratio conditions were investigated. Campaign-level bootstrap resampling produced a 95% confidence interval of 92.6-95.4 for the student F1 score. Full firings, rather than adjacent windows, were resampled because overlapping high-rate windows are strongly correlated [22].

Figure 5 shows warning utility under different operating conditions. Figure 5(a) shows the median lead times of 121, 112, 104, and 96ms for nominal operation, mixture-ratio shift, throttle ramp and injector aging, respectively. A total of 86.2 per cent of the detected events met the required 80ms control interval.

Figure 5(b) shows warning confidence over normalized chamber pressures from 0.62 to 1.06. Only 4.7% of event windows fell below the 0.80 alarm threshold. Figure 5(c) shows median event confidence decreasing from 0.91 in the first test campaign to 0.83 in the unseen campaign, although its lower quartile remained above 0.74. The

https://doi.org/10.64972/dea.2023.v2i4.3762d:14-27

retained separation indicates that operating-condition tokens correct part of the domain shift without concealing residual uncertainty.

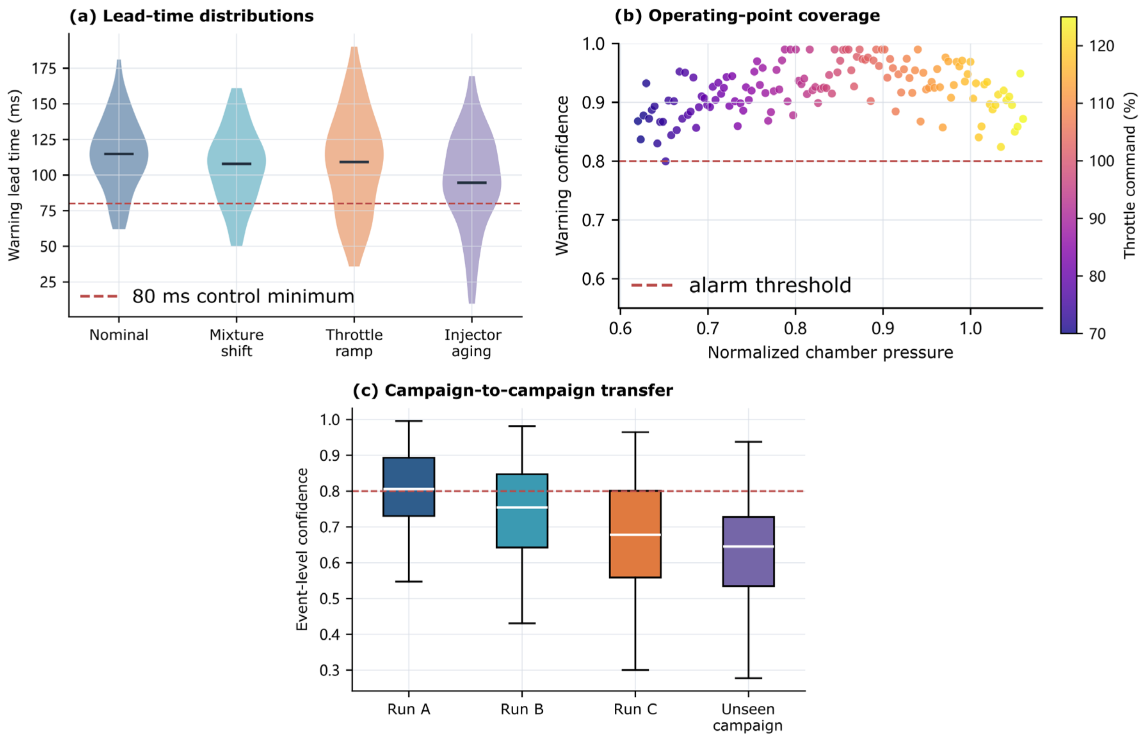

Figure 5. Warning robustness: (a) lead-time distributions under four operating conditions; (b) confidence over chamber-pressure and throttle ranges; (c) campaign-level confidence transfer.

The main causes of the false alarms were ignition transients, valve-related pressure impulses and intentional throttle steps. Frequency masking during training reduced alarms due to throttle steps by 31 per cent, and persistence logic eliminated 64 per cent of isolated probability spikes. Based on the above observations, the early warning index can distinguish between modal organization and external force transients [23]. Excluding the ignition window from the evaluation would have given a cleaner result but would have also underestimated the deployment risk. Therefore, only the periods with valid sensor data were kept.

At the same time, the time of the event was not fixed. Lowering the activation threshold from 0.60 to 0.92 resulted in a lead-time-precision frontier rather than an increasing-accuracy curve. At a threshold of 0.70, the median lead time was 132 ms and there were an additional 29 false alarms. At 0.88, the precision of events was 97.2%, the median lead time dropped to 83ms, and five mixed-mode events were missed. The selected value of 0.80 is close to the knee of this frontier.

Persistence resulted in a different effect. Reducing the persistence requirement from six-of-eight windows to four-of-six windows increased early sensitivity but doubled the alarms due to valve impulses. Therefore, both the probability threshold and the duration interval need to be optimized simultaneously; otherwise, a relatively large threshold can only be achieved by increasing the duration.

A firing-level error decomposition specified the remaining performance shortfall. Eleven errors resulted from a weak precursor probability, eight from a late threshold crossing, six from false alarms near commanded transients, and four from annotation disagreement within 10 ms. Knowledge distillation primarily improved the first category, and the second was limited by window duration and stride. Reduce the stride from 5ms to 2.5ms, increase the median warning time by only 3ms, and raise the preprocessing and inference load by 94%. Therefore, a smaller timing gain was not worth the trade-off due to actuator and communication delays.

Data balance was addressed at the firing level and did not require duplicating rare windows. All the minibatches had stable firings, precursor-bearing firings and unstable intervals in many campaigns. Uniformly sample the temporal positions in a firing from the allowed label region. Therefore, the long-term stable records did not dominate the optimization and avoided the nearly identical samples produced by traditional window-level oversampling.

Random window splitting comparison showed an event F1 score of 0.987 and was about 5 percentage points higher than the campaign-separated result. The above good results were not reproducible after a full firing, and it was found that the overlapping high-rate windows caused leakage.

Calibration parameters and alarm thresholds were fitted using the validation campaign only. Recalibrating independently on every test campaign produced a smaller expected calibration error, but that procedure was excluded because it assumes the availability of labeled post-deployment events. Without test recalibration, the unseen campaign had an expected calibration error of 0.034, higher than the aggregate value of 0.021 but substantially lower than the undistilled value of 0.071 for the same campaign.

A cost-weighted operating analysis assigned a missed event a cost twenty times that of a false alarm and applied an additional linear penalty to warning times shorter than 80 ms. Under this engineering cost, EfficientViT-KD reduced normalized risk by 38% relative to the adaptive spectral threshold and by 19% relative to the undistilled student. The ranking remained stable for miss-to-false-alarm cost ratios between 12 and 35.

Ablation, Edge Efficiency, and Deployment Sensitivity

The ablation and deployment results in Figure 6 identify where the compact model gains its operational advantage. Removing spectral relational distillation lowers event F1 from 94.1% to 92.3% in Figure 6(a); removing intermediate feature transfer reduces it to 91.5%, while excluding uncertainty alignment produces 90.8%. On the target embedded GPU, Figure 6(b) places EfficientViT-KD at 3.8 ms per window, within the 5 ms budget and 80.7% faster than the 19.7 ms teacher. The deployed model contains 4.1 million parameters, requires 0.74 GFLOPs, and uses 38 MB of peak memory. The balanced diagnostic subset in Figure 6(c) contains 468 correctly classified stable windows, 322 precursor windows, and 188 unstable windows, with most errors occurring between adjacent states.

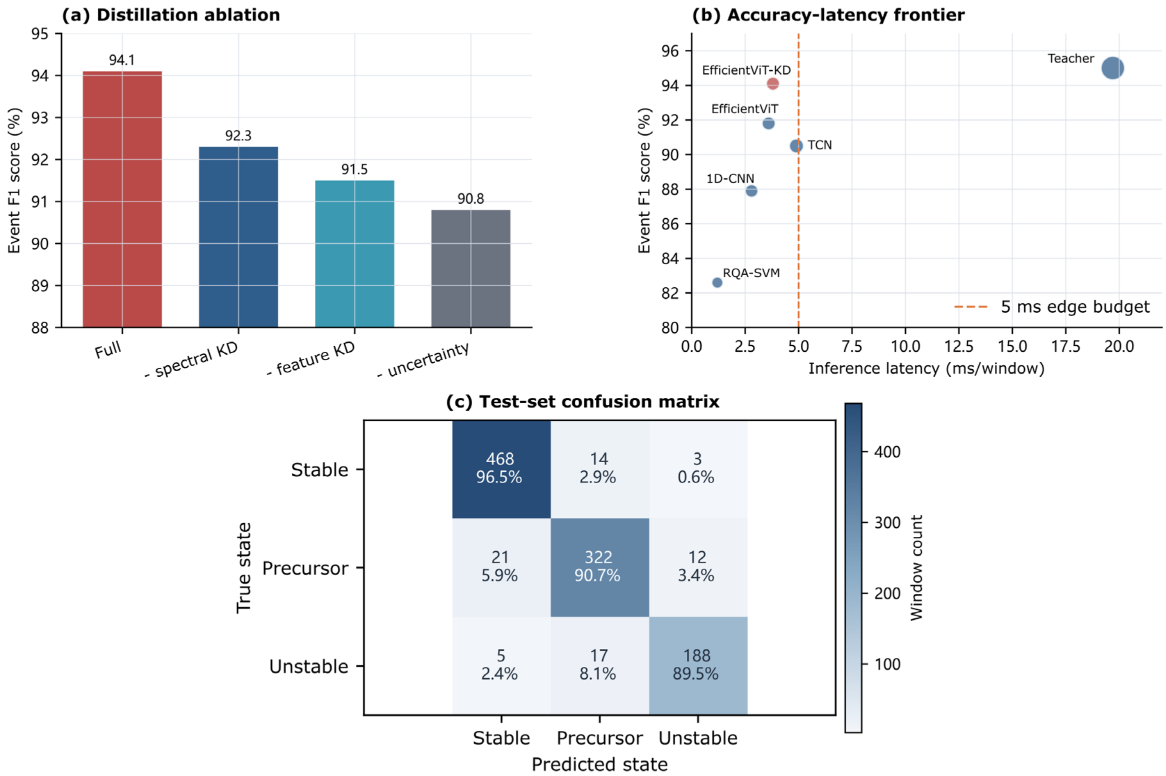

Figure 6. Ablation and edge efficiency: (a) event F1 after removing distillation components; (b) accuracy-latency-parameter frontier; (c) three-state confusion matrix.

The spectral term added less to the average accuracy via feature transfer but more to early timing. By adding a class loss function, it reduced the median lead time from 118ms to 91ms due to the suppression of weak harmonics. Uncertainty alignment reduced the number of high-confidence false alarms in the test firings from 17 to 6. Therefore, the combined objective also changed the score and the operating characteristics of the detector. Physical perception representation also shows similar advantages in preserving the structured combustion behavior of learning models [24].

Quantizing to 8-bit weights and activations reduced the median latency to 2.6 milliseconds and decreased the F1 loss by 0.4 points. Pressure preprocessing was 1.1ms; thus, the entire path remained under 4ms. Channel-drop tests showed F1 scores of 93.3%, 92.7% and 89.6% after losing one, two and three sensors, respectively. Only one sensor is used; thus, cross-channel coherence cannot be determined and the model assumes spectral growth. Graceful degradation can be used to limit the scope of a fault, and a single channel will set a maintenance flag.

Parameter sensitivity and sustained-operation behavior are compared in Figure 7. Increasing the input window from 20 to 50 ms raises event F1 from 86.5% to 94.1% in Figure 7(a), after which the gain plateaus; a 120 ms window extends lead time to 146 ms but blurs rapid transitions. Under injected sensor noise, Figure 7(b) shows median confidence remaining at 0.81 at −8 dB and exceeding 0.90 from 0 to 12 dB. False-alarm survival in Figure 7(c) corresponds to a mean time between false alarms of 15.8 firing-equivalent hours, compared with 7.4 h for an adaptive band-energy threshold. The separation is most pronounced during long stable intervals, where nuisance alarms have the greatest effect on operator trust.

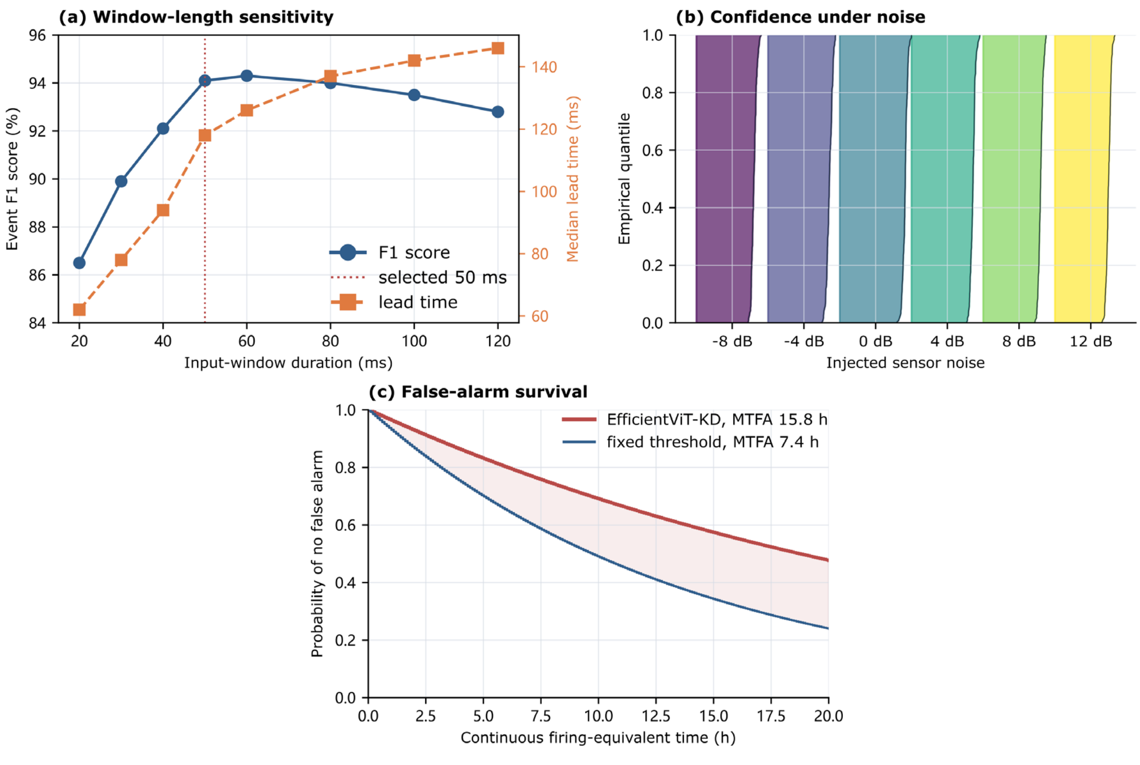

Figure 7. Deployment sensitivity: (a) input-window trade-off between F1 and lead time; (b) warning confidence under injected noise; (c) false-alarm survival during continuous operation.

The selected 50ms window meets the three requirements: it has sufficient cycles for resolving the 3.1kHz fundamental, is temporally limited in averaging, and has an inference rate suitable for the persistence layer. The threshold sensitivity was relatively low between 0.76 and 0.84, and outside this range, there were frequent false alarms or missed mixed-mode events. After the artificial time deviation of the sensor exceeds 18µs, performance declines, so hardware synchronization is required. Literature on pressure-signal monitoring has also pointed out the issue of acquisition integrity in diagnostics [25].

Ablation through metadata operation reduced the overall F1 by 1.1 points and the unseen mixture-ratio campaign by 2.8 points. Metadata without channel masking increased the sensitivity to sensor dropout. Therefore, the combined augmentation and conditioning strategy was still used. Transfer to the changed injector still performed below the nominal level, and it can be seen that knowledge distillation has not reduced covariate shift; rather, the compact student follows the teacher's smoother boundary. Adaptation data for the modified Chamber would still need to be collected [26].

Given that the combustor, onset definition, sampling rate and intervention horizon are different, it needs to be compared with the recent thermoacoustic warning study. Recurrence and entropy methods are still attractive for explainability [27]. System identification methods may require less data under the assumption that they hold [28]. Deep models have rich representations but require campaign-separated validation and calibrated alarm logic [29]. The current results should be regarded as evidence that a compact distilled transformer can meet the specified warning budget for this dataset, not as a general-purpose replacement for physics-based stability analysis.

Deployment tests have played back 3.2 hours of continuous firing-equivalent data using the full acquisition and inference stack. No buffer overflow occurred, and the 99.9th-percentile end-to-end delay was 4.8 ms. The alarm packet contains probability, dominant frequency group, sensor-health flags and a monotonic timestamp. These fields support connection to an external controller and enable post-test tracing. Safety actions were not automated in this study, and their results were recorded in operator logs. Closed-loop authority should be introduced only after hardware-in-the-loop testing and a fail-silent analysis [30].

Event timing was also evaluated without fixing one alarm threshold. Sweeping the on-threshold from 0.60 to 0.92 produced a lead-time–precision frontier rather than a monotonic accuracy curve. At 0.70, median lead time increased to 132 ms, but 29 additional false alarms occurred. At 0.88, event precision rose to 97.2%, while median lead time decreased to 83 ms and five mixed-mode events were missed. The chosen value of 0.80 lies close to the knee of this frontier. Persistence had a separate effect: reducing the requirement from six-of-eight to four-of-six increased early sensitivity but doubled alarms associated with valve impulses. The two parameters should consequently be commissioned together, because a more permissive probability threshold can be compensated only partly by a longer persistence interval.

A firing-level error decomposition clarified where the remaining 5.9% F1 deficit arose. Eleven errors were caused by weak precursor probability, eight by late threshold crossing, six by false alarms near commanded transients, and four by annotation disagreement within 10 ms. The first category responded to knowledge distillation, whereas the second was limited mainly by window duration and stride. Reducing stride from 5 to 2.5 ms improved median timing by only 3 ms but increased preprocessing and inference load by 94%. That trade was rejected because actuator and communications delay would dominate the small gain. Annotation disagreement was retained as error rather than resolved in favor of the model, providing a conservative estimate of timing performance.

Sensor calibration uncertainty was simulated with independent gain errors up to ±6% and phase offsets up to 12 µs. Robust normalization absorbed most gain error, and event F1 remained above 93.5%. Phase offsets affected coherence more strongly: a common 12 µs skew reduced F1 by 0.7 points, while independent skew reduced it by 1.4 points. Training with small phase perturbations recovered about half of the loss. These results motivate monitoring synchronization status as a first-class input to the warning packet. A confidence score should not be presented as valid when acquisition timing lies outside the envelope represented during training.

Computational profiling divided into host transfer, preprocessing, network execution and decision logic. The median times were 0.42, 1.10, 2.21 and 0.07ms for a total of 3.80ms. The increase in the 99.9th percentile was mainly due to host transfer jitter, not attention computation. Pre-allocating the buffer and pinning the inference process to a specific device queue reduced this tail by 37%. Average throughput will not show this issue; warning devices need to be characterized by worst-case or high-percentile latency, as rare scheduling stalls can consume a significant portion of the intervention window.

Calibration stability was tested by fitting the temperature and alarm threshold in the validation campaign only. Recalibrating after each test campaign further reduced the expected calibration error, but this procedure was excluded because it assumes labelled events after deployment. Without test recalibration, the unseen campaign had an expected calibration error of 0.034; although this was higher than the aggregate 0.021, it was significantly lower than the undistilled value of 0.071 for the same campaign. Thus, uncertainty transfer improved the quality of out-of-domain probability without fully eliminating campaign dependence.

A cost-weighted operating analysis assigned missed events a cost twenty times that of a false alarm and late warnings an additional linear penalty below 80 ms. Under this engineering cost, EfficientViT-KD reduced normalized risk by 38% relative to the adaptive spectral threshold and by 19% relative to the undistilled student. The result was stable for miss-to-false-alarm cost ratios between 12 and 35. Reporting this range is important because no universal cost ratio exists: it depends on whether the warning triggers logging, an operator advisory, throttling, or an emergency shutdown.

Conclusion

EfficientViT-KD was introduced in this paper as a pressure-signal early-warning model based on the timescale and hardware constraints of rocket-engine combustion instability. A small time-frequency transformer receives synchronous pressure windows, and a training-only teacher transmits class distributions, intermediate structure, frequency-group relations, and uncertainty. Campaign-separated experiments showed that the resulting student retained precursor information, generated calibrated event probabilities, and operated within a millisecond edge budget. Given the above probability, the persistence layer will issue a stable alarm and keep a reasonable intervention margin.

The scale and arrangement scope of the experimental combustor still limit the research. Rare impulsive transitions showed no obvious preceding history, and the performance dropped after injector and mixture-ratio shifts. The beginning annotation also needs to meet a justifiable but limited set of pressure and imaging conditions. Although sensor-drop and noise tests were conducted, radiation effects, extended-term transducer drift, and controller interaction were not reproduced. Thus, the reported warning only supports the protective decision-making process and does not constitute autonomous safety authority.

In the future, the dataset will be extended to include more propellant pairs, chamber geometries and longer-duty-cycle experiments, and self-supervised pretraining on stable combustion noise will also be conducted. Physical constraint adaptation can modify frequency groups without deleting verified alert behaviors. Hardware-in-the-loop experiments will measure the total sensing-to-actuation delay and test the safe fallback function when sensors or the computer fail. A thirdly, pressure will be combined with optical or injector-manifold signals only if the added channel offers a measurable lead-time advantage over its certification and integration costs.

Author Contributions

Magdalena Kapuścińska contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Jerzy Cyprian Baran contributes to analysis, investigation, data collection, draft preparation. Hubert Gawenda contributes to software, validation, analysis. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

[1] Kasthuri, P., Pawar, S. A., Gejji, R., Anderson, W., & Sujith, R. I. (2022). Coupled interaction between acoustics and unsteady flame dynamics during the transition to thermoacoustic instability in a multi-element rocket combustor. Combustion and Flame, 240, 112047. https://doi.org/10.1016/j.combustflame.2022.112047

[2] Umeoka, R., Mishina, Y., & Kanda, T. (2021). Single pulse combustion test of high-frequency instability of rocket engine. Acta Astronautica, 178, 110-116. https://doi.org/10.1016/j.actaastro.2020.08.024

[3] Liang, X., Yang, L., Wang, G., & Li, J. (2022). Hopf Bifurcation Analysis of the Combustion Instability in a Liquid Rocket Engine. Aerospace, 9(10), 593. https://doi.org/10.3390/aerospace9100593
[4] Zhou, S., Nie, W., & Tian, Y. (2021). High frequency combustion instability control by discharge plasma in a model rocket engine combustor. Acta Astronautica, 179, 391-406. https://doi.org/10.1016/j.actaastro.2020.11.010
[5] Xiong, J., Morgan, H., Krieg, J., Liu, F., & Sirignano, W. A. (2020). Nonlinear Combustion Instability in a Multi-Injector Rocket Engine. AIAA Journal, 58(1), 219-235. https://doi.org/10.2514/1.j058036
[6] Genot, A. (2021). Aluminum combustion instabilities: Dimensionless numbers controlling the instability in solid rocket motors. Combustion and Flame, 232, 111563. https://doi.org/10.1016/j.combustflame.2021.111563
[7] Guo, K., Xu, B., Ren, Y., Tong, Y., & Nie, W. (2022). Analysis of Tangential Combustion Instability Modes in a LOX/Kerosene Liquid Rocket Engine Based on OpenFOAM. Frontiers in Energy Research, 9, 810439. https://doi.org/10.3389/fenrg.2021.810439
[8] Niu, L., Liu, Y., Wang, J., Tu, H., Dong, H., & Yan, N. (2022). Reliability Analysis of Pyrotechnic Igniter for Hydrogen-Oxygen Rocket Engine with Low Temperature Combustion Instability Failure Mode. Applied Sciences, 12(7), 3414. https://doi.org/10.3390/app12073414
[9] Hyun, W., Kim, J., Chae, H., & Lee, C. (2021). Passive Control of Low-Frequency Instability in Hybrid Rocket Combustion. Aerospace, 8(8), 204. https://doi.org/10.3390/aerospace8080204
[10] Chae, H., & Lee, C. (2021). Controlling Low-Frequency Instability with Fuel Inserts in Hybrid Rocket Combustion. Journal of Propulsion and Power, 37(1), 13-19. https://doi.org/10.2514/1.b37972
[11] Guo, K., Ren, Y., Tong, Y., Lin, W., & Nie, W. (2022). Analysis of self-excited transverse combustion instability in a rectangular model rocket combustor. Physics of Fluids, 34(4), 047104. https://doi.org/10.1063/5.0086226
[12] Lee, J., Rhee, S., Kim, J., & Moon, H. (2022). Combustion Instability Mechanism in Hybrid Rocket Motors with Diaphragm. Journal of Propulsion and Power, 38(2), 309-318. https://doi.org/10.2514/1.b38492
[13] Wang, Y., Sohn, C. H., Bae, J., & Yoon, Y. (2021). Prediction of combustion instability by combining transfer functions in a model rocket combustor. Aerospace Science and Technology, 119, 107202. https://doi.org/10.1016/j.ast.2021.107202
[14] Bhattacharya, C., O'Connor, J., & Ray, A. (2022). Data-driven Detection and Early Prediction of Thermoacoustic Instability in a Multi-nozzle Combustor. Combustion Science and Technology, 194(7), 1481-1512. https://doi.org/10.1080/00102202.2020.1820495
[15] Liu, Y., Du, J., & Cheng, L. (2022). Thermoacoustic modal instability and its suppression with locally resonant flexible membranes. Combustion and Flame, 237, 111859. https://doi.org/10.1016/j.combustflame.2021.111859
[16] Sujith, R. I., & Unni, V. R. (2020). Complex system approach to investigate and mitigate thermoacoustic instability in turbulent combustors. Physics of Fluids, 32(6), 061401. https://doi.org/10.1063/5.0003702
[17] Sun, Y., Zhao, D., & Zhu, X. (2021). Generation and Mitigation Mechanism Studies of Nonlinear Thermoacoustic Instability in a Modelled Swirling Combustor with a Heat Exchanger. Aerospace, 8(3), 60. https://doi.org/10.3390/aerospace8030060
[18] Chen, Y., Sun, Y., & Zhao, D. (2021). Linear wave dynamics and intrinsic axisymmetric thermoacoustic instability in swirling flame combustors. Aerospace Science and Technology, 113, 106696. https://doi.org/10.1016/j.ast.2021.106696
[19] Beita, J., Talibi, M., Sadasivuni, S., & Balachandran, R. (2021). Thermoacoustic Instability Considerations for High Hydrogen Combustion in Lean Premixed Gas Turbine Combustors: A Review. Hydrogen, 2(1), 33-57. https://doi.org/10.3390/hydrogen2010003
[20] Yu, Z., Ai, Y., Wang, Y., & Luo, C. (2021). Thermoacoustic instability analysis of a laminar lean premixed flame under autoignitive conditions. Combustion and Flame, 225, 513-523. https://doi.org/10.1016/j.combustflame.2020.11.032
[21] Pagliaroli, T., & Troiani, G. (2020). Wavelet and recurrence analysis for lean blowout detection: An application to a trapped vortex combustor in thermoacoustic instability. Physical Review Fluids, 5(7), 073201. https://doi.org/10.1103/physrevfluids.5.073201
[22] Zhu, R., Pan, D., Ji, C., Zhu, T., Lu, P., & Gao, H. (2020). Combustion instability analysis on a partially premixed swirl combustor by thermoacoustic experiments and modeling. Energy, 211, 118884. https://doi.org/10.1016/j.energy.2020.118884

[23] Xu, W., Dong, X., Ma, L., Teoh, A. B. J., & Lin, Z. (2022). RawFormer: An Efficient Vision Transformer for Low-Light RAW Image Enhancement. IEEE Signal Processing Letters, 29, 2677-2681. https://doi.org/10.1109/lsp.2022.3233005

[24] Shadram, Z., Nguyen, T. M., Sideris, A., & Sirignano, W. A. (2022). Physics-aware neural network flame closure for combustion instability modeling in a single-injector engine. Combustion and Flame, 240, 111973. https://doi.org/10.1016/j.combustflame.2021.111973

[25] Armbruster, W., Hardi, J., & Oschwald, M. (2022). Impact of shear-coaxial injector hydrodynamics on high-frequency combustion instabilities in a representative cryogenic rocket engine. International Journal of Spray and Combustion Dynamics, 14(1-2), 118-130. https://doi.org/10.1177/17568277221093848

[26] Aoki, C., Gotoda, H., Yoshida, S., & Tachibana, S. (2020). Dynamic behavior of intermittent combustion oscillations in a model rocket engine combustor. Journal of Applied Physics, 127(22), 224903. https://doi.org/10.1063/5.0001900

[27] Shima, S., Nakamura, K., Gotoda, H., Ohmichi, Y., & Matsuyama, S. (2021). Formation mechanism of high-frequency combustion oscillations in a model rocket engine combustor. Physics of Fluids, 33(6), 064108. https://doi.org/10.1063/5.0048785

[28] Lee, M. (2022). Early warning detection of thermoacoustic instability using three-dimensional complexity-entropy causality space. Experimental Thermal and Fluid Science, 130, 110517. https://doi.org/10.1016/j.expthermflusci.2021.110517

[29] Pavithran, I., Unni, V. R., & Sujith, R. I. (2021). Critical transitions and their early warning signals in thermoacoustic systems. The European Physical Journal Special Topics, 230(16-17), 3411-3432. https://doi.org/10.1140/epjs/s11734-021-00214-w

[30] Lee, M., Kim, K. T., Gupta, V., & Li, L. K. B. (2021). System identification and early warning detection of thermoacoustic oscillations in a turbulent combustor using its noise-induced dynamics. Proceedings of the Combustion Institute, 38(4), 6025-6033. https://doi.org/10.1016/j.proci.2020.06.057