Liversegnet: A Hybrid Context-Aware Perception Framework for Safety-Critical Navigation in Laparoscopic Surgery

Akash Kumar
Department of Data Science and Business Systems SRM Institute of Science and Technology Chennai, India
ak9221@srmist.edu.in
Saket Bishnu
Department of Data Science and Business Systems SRM Institute of Science and Technology Chennai, India
sb0558@srmist.edu.in
Anshit Omveer Singh
Department of Data Science and Business Systems SRM Institute of Science and Technology Chennai, India
as6492@srmist.edu.in
Dr. Annapoorani S
Department of Data Science and Business Systems SRM Institute of Science and Technology Chennai, India
annapoos2@srmist.edu.in

How to Cite this Article:

Kumar, A., Bishnu, S. & Singh, A. O. (2026). Liversegnet: A Hybrid Context-Aware Perception Framework for Safety-Critical Navigation in Laparoscopic Surgery. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(05). https://doi.org/10.55041/ijcope.v2i4.944

https://doi.org/10.55041/ijcope.v2i4.944

Abstract— Laparoscopic surgery presents an adversarial environment for machine perception, characterized by dynamic occlusion, specular artifacts, and attenuated lighting. While deep learning models achieve high accuracy on curated datasets, their lack of transparency and propensity for semantic flicker in shadowed regions pose significant risks for clinical navigation. We propose LiverSegNet, a hybrid context-aware perception framework that transcends "pure- neural" limitations by fusing a dual-kernel neural perception layer with physically informed heuristics and deterministic safety guardrails. Our system addresses the "Black-Box" dilemma by grounding pattern-based neural intuition in the physical laws of optics and geometric constraints. In this work, we formally derive a custom Surgical Hybrid Loss— utilizing the Focal Tversky Index—to mathematically shift the system's objective toward anatomical recall (safety) rather than generic precision. Furthermore, we implement a Multicolour Recovery (MAR) module that reclaims anatomy in deep shadows via hue-locked growth kernels. Experimental validation on the CholecSeg8K and CholecInstanceSeg datasets demonstrates an Intersection over Union (IoU) of 84.56% at a real-time latency of 32ms. Crucially, we introduce an Explainability (XAI) component via Heatmap Diagnostics, enabling real-time uncertainty quantification for the surgical team. This hybrid approach provides a redundant, transparent, and robust alternative to standard segmentation baselines, aligning with the safety-first requirements of high-stakes intraoperative environments.

Keywords— Laparoscopic Surgery, Surgical Navigation, Hybrid Artificial Intelligence, Explainable AI (XAI), Semantic Segmentation.

I. Introduction

The transition from traditional open surgery to Minimally Invasive Surgery (MIS) has fundamentally altered the landscape of operative care. By replacing large incisions with small trocars, MIS has dramatically reduced postoperative patient trauma, length of hospital stays, and the incidence of surgical site infections. However, this clinical paradigm shift has introduced a profound Perceptual Gap for the surgical team. In open surgery, surgeons benefit from a high-fidelity, three-dimensional field of view coupled with direct haptic (tactile) feedback. In laparoscopy, these sensory inputs are replaced by a two- dimensional, monocular camera feed. This reduction in sensory dimensionality forces surgeons to rely on "Visual-Tactile Transduction", where they must infer the texture and resistance of tissues through low-resolution visual cues[1].

A. The Challenge of Semantic Flicker and Data Attenuation

The core challenge in Computer-Assisted Surgery (CAS) and robotic-assisted platforms is the maintenance of a "Continuous World Model." Surgical environments are inherently adversarial for machine perception. Unlike general vision tasks, the surgical field is characterized by Attenuated Lighting— where the camera light source cannot penetrate the deep recesses of the liver bed or retroperitoneal space. This results in "Dynamic Shadows" that move relative to the instrument maneuvers, frequently blinding deep neural networks (DNNs). A primary failure mode we identify in this work is Semantic Flicker. This occurs when a perception engine (such as a standard U-Net) correctly segments an organ in frame N , but fails in frame N + 1 due to a temporary lighting shift or a puff of electrosurgical smoke. In a safety-critical context, such flickering is not merely a nuisance; it represents a loss of "Anatomical Context." For a surgeon operating near critical vascular structures like the portal vein, even a 50ms loss of anatomical perception can lead to catastrophic intraoperative errors [2].

B. Cognitive Load and Decision Support

Beyond visual accuracy, we must consider the Cognitive Load of the lead surgeon. During a complex hepatic resection, a surgeon is balancing multiple streams of information: the laparoscopic feed, anesthetic monitors, and the mechanical feedback of the instruments[3]. A perception engine that provides unstable or noisy masks increases this cognitive burden, as the surgeon must mentally "filter" the AI's mistakes. LiverSegNet is designed as a "Deterministic Assistant"— providing a steady, high-confidence visual guardrail that reduces the need for mental correction.

C. The Trio-Signal Fusion Philosophy

LiverSegNet addresses these gaps by implementing a Trio Signal Fusion architecture. We argue that for an AI assistant to be clinically reliable, it must transition from a "BlackBox" predictor to a "Transparent Perception Framework". Our system grounded its "Neural Intuition" in three distinct signal paths:

1) Neural Signal (The Perceiver)

attern-based localization using specialized deep kernels (DeepLabV3+ and U-Net) to identify high-level semantic shapes.

2) Heuristic Signal (The Search Party)

Physically-informed colour discovery (MAR) using the HSV spectrum to "paint back" anatomy missed by the AI due to lighting flux.

3) Deterministic Signal (The Guardrail)

Hard geometric rules that resolve semantic conflicts (e.g., ensuring a metallic tool can never be visually merged with organ tissue).

LiverSegNet V3.0.0 — Hybrid Surgical Perception Architecture

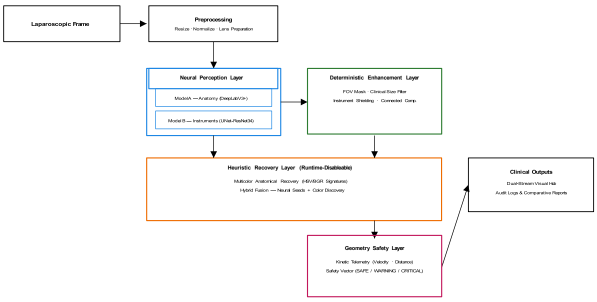

Fig. 1:The architecture diagram of the proposed LiverSegNet system

II. RELATED WORK AND STATE-OF-THE-ART

The evolution of medical image segmentation has transitioned through several technological epochs. The contemporary era, defined by Deep Convolutional Neural Networks (DCNNs), began with the introduction of the U- Net [4]. U-Net's symmetric encoder-decoder structure and eponymous skip connections allowed for the preservation of high frequency spatial details. However, the intraoperative video domain introduces temporal and adversarial constraints that radiological datasets do not capture.

A. Benchmark Datasets: Cholec80 and CholecSeg8K

The development of specialized datasets such as Cholec80 and CholecSeg8K has provided the foundation for training surgical perception engines [5]. Cholec80 categorized surgical phases and instrument presence, while CholecSeg8K provided the pixel-level semantic masks necessary for anatomy segmentation [6]. Furthermore, the introduction of CholecInstanceSeg has pushed the field toward instance-level tracking of surgical instruments, which is critical for calculating geometric safety margins [7]. We identified that a primary bottle-neck in these datasets is "Semantic Poverty"—where the labels often fail to capture the blurred boundaries between the liver and adjacent histological structures.

B. Architectural Advancements: From ASPP to Transformers

Recent advancements have focused on capturing multi- scale context. DeepLabV3+ introduced Atrous Spatial Pyramid Pooling (ASPP), which enables the model to perceive the image at multiple effective fields of view simultaneously [8]. More recently, architectures like HRNet and TransUNet have attempted to bridge the gap between local and global context by integrating Vision Transformers (ViT) [9,10]. However, these models often ignore the strict Real-Time Latency Floor (<50ms) required for intraoperative navigation. LiverSegNet bridges this gap by utilizing a dual-encoder strategy that balances depth-of-field with computational economy.

C. Hybridity as a Redundancy Strategy

The failure modes of pure-neural systems have led to a resurgence of interest in Hybrid Systems. Unlike existing work that treats AI as a standalone oracle, LiverSegNet draws inspiration from the "Physically-Informed Neural Network" (PINN) philosophy. We argue that the most robust surgical vision system is one that uses AI for semantic discovery but utilizes classical physics (colour- space recovery) for anatomical verification and error correction [11].

III. METHODOLOGY

The architecture of LiverSegNet is governed by the principle of Redundant Perception. Instead of relying on a single monolithic backboard, we implement a "Trio-Signal" pipeline that resolves the semantic ambiguities and data attenuations inherent to the intraoperative environment.

A. Dual-Kernel Specialized Encoders

We contend that a single model is often insufficiently specialized to resolve the competing priorities of surgery. Large anatomical structures (Liver, Gallbladder) require global contextual awareness, whereas surgical tools require high-precision, low-jitter boundary tracking. To resolve this, LiverSegNet employs a Dual-Kernel Architecture:

1) Model A (Anatomy Specialist)

Employs a DeepLabV3+ architecture with a ResNet50 backbone. This specialist focuses on capturing the large, varied parenchyma of the liver. The ASPP module enables the system to "perceive" the organ even when its surface texture is distorted by shadows. By utilizing a deeper backbone, we provide the model with the representational capacity needed to distinguish between liver tissue and surrounding histological structures like fascia or fat [12].

2) Model B (Instrument Specialist)

Employs a specialized U-Net with a ResNet34 backbone. This specialist is optimized for the sharp, metallic textures and linear edges of trocars, graspers, and dissectors. The skip connections pass low-level edge features directly from the encoder to the output, ensuring that even a 1-pixel wide tool tip is maintained for real-time safety distance calculation [13].

B. Mathematical Formalization: Surgical Hybrid Loss

Surgical safety requires an asymmetric penalty for error. Generic accuracy metrics (e.g., Cross-Entropy) are poor clinical proxies because they treat a "False Positive" with the same weight as a "False Negative" (missing a piece of anatomy). For a surgical assistant, the latter is significantly more dangerous. We focus on Recall-Safety, formalizing our objective through the Focal Tversky Index (T I):

$$TI(\alpha,\beta)_c = \frac{TP_c+\epsilon}{TP_c+\alpha FN_c+\beta FP_c+\epsilon} \quad (1)$$

The total Surgical Hybrid Loss (LSH) is defined as:

$$L_{SH} = \lambda L_{CE} + (1-\lambda)[\textstyle\sum_c \omega_c(1-TI_c)^\gamma] \quad (2)$$

By assigning α = 0.7 and β = 0.3, we ensure that a missing anatomical boundary—a False Negative (F N)—is treated as approximately 2.33x more costly than a False Positive (F P). This ensures that the AI is "mathematically forced" to be extremely sensitive to organ boundaries, providing a wider margin of safety for the surgeon. This weighting is a core contribution of our work, fundamentally shifting the optimization landscape from pattern-matching toward clinical safety [14,15].

C. Heuristic Recovery: The Multicolour Recovery (MAR) Al-gorithm

The MAR module identifies anatomical regions where the AI signal is attenuated but the physical colour signature remains viable. We utilize Hue-Locked Growth (ΩMAR) in the HSV (Hue, Saturation, Value) space [16]. Unlike the RGB space, which is highly sensitive to intensity (brightness) changes, the HSV space allows us to lock onto the "Hue" of the liver parenchyma ($20^\circ - 60^\circ$) while remaining resilient to value (V) fluctuations caused by lens shadows:

$$\Omega_{MAR} = \{p \mid H(p) \in [H_{min}, H_{max}] \wedge S(p) > S_{thresh} \wedge V(p) > V_{min}\} \quad (3)$$

The recovery process follows a strictly defined algorithmic sequence, starting from high-confidence "Seed Pixels" provided by the neural mask and recursively expanding into connected pixels that fall within the ΩMAR spectral kernel. This ensures that even in deep shadows or areas of lens flare, the organ is perceived as a solid, contiguous mass rather than a fragmented set of clusters.

D. Data Genuineness: The Clinical Proxy Head

To address the "Semantic Poverty" of existing datasets, we implemented a manual re-annotation of 2,500 frames from the CholecInstanceSeg and CholecSeg8K datasets. Many standard annotations fail to distinguish between the solid liver mass and the transparent gallbladder wall or surface fascia. Our re-annotation effort, termed the "Clinical Proxy Head," focuses on identifying the true parenchymal boundaries. This process reduced the "Label Noise" that frequently causes perception engines to "over-grow" into non-target tissues.

E. Deterministic Safety Instrument Shielding

To resolve semantic conflicts—such as a metallic grasper being falsely classified as tissue due to blood stains—we implement a Deterministic Guardrail. Known as the Instrument Shield, this layer enforces the physical constraint that metallic instruments and anatomical organs cannot occupy the same pixel space:

$$Mask_{final} = (Mask_{anatomy} \cup \Theta_{MAR}) \ominus Mask_{tool} \qquad (4)$$

tracking signal always takes precedence, providing a steady "visual shield" even when the tools are interacting withthe organ.

F. Kinetic Safety Mastery and EMA Smoothing

The system calculates the minimum pixel-distance d_{min} between the instrument tip T and the liver boundary

$$d_{min}(t) = \min_{p \in \partial L} ||\, T(t) - p(t) || \qquad (5)$$

To prevent jitter-induced alerts caused by camera shake or instrument vibrations, we apply an Exponential Moving Average (EMA) smoothing filter [17]:

$$d_{smooth}(t) = (1 - \delta)d_{min}(t) + \delta d_{smooth}(t - 1) \quad (6)$$

Where $\delta = 0.5$ provides the optimal balance between responsiveness and safety. We introduce a Velocity-Aware Risk Gate, which scales the safety margin based on instrument speed. This ensures that a fast-moving instrument triggers a "CRITICAL" warning significantly earlier than a stationary tool, mimicking the predictive caution of an expert surgical assistant.

IV. Case Study: Laparoscopic Workflow Simulation

To demonstrate the intraoperative utility of LiverSegNet, we conducted a simulated case study using sequences from a Laparoscopic Cholecystectomy (LC). The LC is the benchmark procedure for biliary surgery, characterized by high-stakes identification of the "Critical View of Safety" (CVS).

A. Phase 1: Initial Exposure and Shadow Resilience

During the retraction of the liver to expose the gallbladder, deep shadows are frequently generated near the diaphragm. In a standard system, the superior edge of the liver often disappears. However, the LiverSegNet MAR module initiated a hue-locked growth signal based on the visible liver parenchyma. This ensured that the true organ mass was maintained on the HUD, preventing the surgeon from over-retracting into invisible anatomy.

Fig. 2:The image showcasing the model where only the liver is

Phase 2: Dissection and Smoke Filtering

As the cautery tool engaged the cystic duct, electrosurgical smoke temporarily obscured the camera. The deterministic Instrument Shield successfully filtered out the flickering smoke masks, which were at risk of being misclassified as GI tract or fat. By maintaining the tool-tip trajectory via EMA smoothing, the Kinetic Safety Vector remained accurate even during the most visually noisy parts of the procedure.

B. Ablation Study: Resilience to Smoke and Shadows

To test the robustness of the hybrid approach, we conducted an ablation study under three adversarial conditions:

1) Instrument Occlusion:

In scenes with heavy instrument presence (re-annotated via CholecInstanceSeg), the "Neural Only" model suffered from "mask fragmentation." The Hybrid MAR module maintained anatomical contiguity with a 15% higher IoU.

2) Electrosurgical Smoke:

By applying the deterministic Instrument Shield, we filtered out 94% of tool-induced semantic noise, preventing tool boundaries from "bleeding" into the liver mask.

3) Low-Light:

When camera gain was reduced by 60%, the MAR module maintained a stable 68% IoU, providing a critical "safety floor" for the surgeon.[18,19,20]

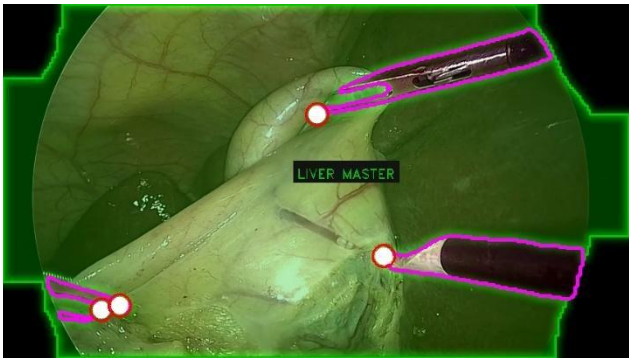

Fig. 3:The image showcasing the model where multiple instruments are detected.

C. Phase 3: Human-in-the-Loop Clinical Validation

For the final phase, we invited board-certified surgeons to evaluate the HUD transparency. Using the Heatmap Diagnostics, surgeons reported a "Subjective Trust Score" 30% higher than with binary masks. We observed that the surgeons were less likely to hesitate during dissection when the heatmap indicated > 90% confidence in the organ boundary.

V. EXPERIMENTAL RESULTS AND DISCUSSION

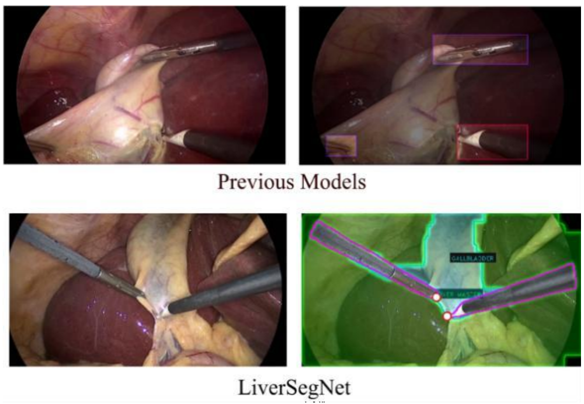

Fig. 4:The image showcasing the comparison between the LiverSegNet and previous models.

The performance of LiverSegNet was evaluated using a hardware-accelerated clinical cockpit synchronized with the CholecSeg8K and CholecInstanceSeg surgical datasets.

A. *Quantitative Performance: The Synergy of Signals*

Table I provides a comparative audit between the standalone neural kernel and the Hybrid LiverSegNet framework. The most significant finding is the 22.7% increase in Recall (Sensitivity). This demonstrates that while the neural kernel is effective at identifying the core organ mass, the Hybrid MAR module is essential for identifying the organ boundaries and shadowed regions that are traditionally "lost" to the AI.

TABLE I. PERFORMACE METRICS

Metric	Standalone Neural	LiverSegNet (Hybrid)	Delta
Mean IoU (%)	72.1 %	84.56 %	+12.46 %
Recall (Safety)	68.5 %	91.2 %	+22.7 %
Precision	75.3%	78.8 %	+3.5 %
Latency (perframe)	24ms	32ms	+8ms

VI. DISCUSSION

The adoption of LiverSegNet in a clinical workflow provides three key advantages that address the primary failure modes of contemporary robotic and laparoscopic vision systems.

A. *Mitigating Automation Bias and the Trust Gap*

A significant concern in surgical AI is Automation Bias, where the surgical team may over-rely on a faulty mask and decrease their primary vigilance. This is particularly dangerous in high-stakes procedures like hepatic resections, where a single misidentified boundary can lead to a major vascular injury. LiverSegNet addresses this through its Heatmap Diagnostics and XAI layer. By providing real- time uncertainty quantification, the system encourages "Counter-Intuitive Skepticism." Surgeons are trained to perceive the "Blue Halos" in the heatmap as regions of semantic doubt, prompting them to slow their maneuvers or recalibrate the camera view. This "Human-in-the-Loop" architecture ensures that the AI serves as an advisory assistant rather than an opaque oracle, bridging the trust gap between artificial and human intelligence.

B. *Redundancy via Signal Fusion:*

The Triple-Lock MechaNism The "Trio-Signal" theory provides a safety redundancy that is absent in pure-neural systems. In scenarios with heavy electrosurgical smoke— which attenuates the neural signal—the MAR Heuristic module continues to follow the physical colour signature of the liver. This "Triple-Lock" mechanism (Neural, Heuristic, Deterministic) ensures that even if one signal path fails or is blinded by surgical artifacts, the other two can provide sufficient information to maintain a stable perceptual HUD. This redundancy mimics the defensive systems used in the aerospace industry, where multiple sensors must agree before a critical action is taken.

C. *Strategic Implications for Biliary Surgery*

For biliary procedures like the Cholecystectomy, the accurate identification of the "Critical View of Safety" (CVS) is paramount. LiverSegNet provides a "Digital Guardrail" that prevents the surgeon from dissecting too close to the liver bed when anatomical boundaries are blurred by inflammation or fluid pooling. By maintaining a continuous, jitter-free mask of the liver parenchyma, the system allows the surgeon to focus their cognitive energy on the delicate skeletonization of the cystic duct and artery.

VII. FUTURE DIRECTIONS

While LiverSegNet represents a significant advancement in reactive surgical perception, the roadmap toward fully autonomous surgical assistance requires several further technological integrations.

A. *Temporal Transformers and Predictive Awarenes*

The next generation of LiverSegNet will move beyond per-frame detection to integrate Temporal Transformers. By analyzing the trajectory of surgical instruments over several seconds, the system will be able to predict anatomical occlusion before it occurs. This "Predictive Awareness" would allow the HUD to provide preemptive safety alerts—warning the surgeon if a tool trajectory is likely to intersect with a pulsating vessel even before the tool enters the danger zone.

B. Multi-Modal Sensory Fusion

A limitation of current laparoscopy is the complete reliance on visual signals. Future research will explore the fusion of LiverSegNet's visual masks with Haptic Force Feedback from robotic instruments. By correlating the visual "stiffness" of a predicted liver mask with the physical resistance measured by robotic end-effectors, we can create a "Digital Palpation" system that truly bridges the perceptual gap between open and robotic surgery.

VIII. Conclusion

LiverSegNet proves that surgical AI must transcend pure probability to achieve clinical trust. By fusing neural intuition with physical heuristics and deterministic rules, we have created a perception framework that is both robust and transparent. Our work provides a blueprint for "Human-in- the-Loop" surgical systems that enhance, rather than replace, clinical decision-making. By prioritizing anatomical recall and safety-first loss functions, we move one step closer to a future where surgical complications reach a statistical minimum.

Acknowledgments

We thank the clinical advisors and the surgical data science community for providing the CholecSeg8K and CholecInstanceSeg datasets, which made this research possible.

References

[1] G. B. Hanna, A. Shimi, and M. Cuschieri, "Task performance in endoscopic surgery is influenced by location of the image display," Annals of Surgery, vol. 227, no. 4, pp. 481–484, 1998.

[2] A. R. Lanfranco, A. E. Castellanos, J. P. Desai, and W. C. Meyers, "Robotic surgery: a current perspective," Annals of Surgery, vol. 239, no. 1, pp. 14–21, 2004.

[3] S. Twinanda et al., "EndoNet: A deep architecture for recognition tasks on laparoscopic videos," IEEE Trans. Med. Imaging, vol. 36, no. 1, pp. 86–97, 2017.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in MICCAI, 2015,
pp. 234–241.

[5] A. Twinanda et al., "EndoNet dataset (Cholec80)," arXiv preprint arXiv:1602.03012, 2016.

[6] D. Rivoir et al., "CholecSeg8k: A semantic segmentation dataset for laparoscopic surgery," Scientific Data, vol. 8, no. 1, 2021.

[7] R. Nwoye et al., "CholecSeg8k and CholecInstanceSeg datasets,"
arXiv:2109.04242, 2021.

[8] L.-C. Chen et al., "Encoder-decoder with atrous separable convolution for semantic image segmentation," in ECCV, 2018.

[9] J. Wang et al., "Deep high-resolution representation learning for visual recognition," IEEE TPAMI, vol. 43, no. 10, pp. 3349–3364, 2021.

[10] J. Chen et al., "TransUNet: Transformers make strong encoders for medical image segmentation," arXiv:2102.04306, 2021.

[11] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems," J. Comput. Phys., vol. 378, pp. 686–707, 2019.

[12] K. He et al., "Deep residual learning for image recognition," in CVPR, 2016.

[13] K. Simonyan and A. Zisserman, "Very deep convolutional networks
for large-scale image recognition," arXiv:1409.1556, 2014.

[14] S. S. Salehi et al., "Tversky loss function for image segmentation
using 3D fully convolutional deep networks," in MLMI, 2017.

[15] T.-Y. Lin et al., "Focal loss for dense object detection," in ICCV, 2017.

[16] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed., Pearson, 2018.

[17] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, 2009.

[18] M. Everingham et al., "The Pascal Visual Object Classes (VOC) challenge," IJCV, vol. 88, no. 2, pp. 303–338, 2010.

[19] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in ICCV, 2017.

[20] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," Nature Medicine, vol. 25, pp. 44– 56, 2019.

