# VISTA: A Benchmark for Real-Time Video Streaming under Network Impairments in Surgical Teleoperation

Zexin Deng*, Zhenhui Yuan*, Tian Lu*, Gaofeng Li†, Meipeng Huang‡, and Longhao Zou§

***Abstract*—Real-time video streaming is crucial in surgical teleoperation, yet reproducible evaluation under realistic network impairments remains limited. This paper presents *VISTA*, a benchmark designed to study how impairments along the forward video path affect received video quality, temporal continuity, and human task performance. VISTA employs Linux Traffic Control with NetEm and a Gilbert-Elliott loss model to emulate five network conditions: Hospital LAN, 5G urban, 4G rural, LEO satellite, and GEO Satellite. The benchmark integrates a standardised peg transfer task with synchronized measurements of Quality of Service (QoS), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and freeze rate, while maintaining a stable reverse control channel. Across 375 experimental trials, network degradation substantially reduced teleoperation performance: success rate decreased from 97% in Hospital LAN to 79% in 5G urban, 35% in 4G rural, 71% in LEO satellite, and 12% in GEO Satellite, while mean task completion time for successful trials increased from 80 s in Hospital LAN to 117 s in 5G Urban, 211 s in 4G Rural, 152 s in LEO Satellite, and 255 s in GEO Satellite. These findings show that network impairments have a direct impact on task completion and success in surgical teleoperation, and provide a reproducible basis for evaluating teleoperation video under realistic network constraints. Source code available: https://github.com/Dzxx623/VISTA.**

***Index Terms*—real-time video streaming, surgical teleoperation, hardware-in-the-loop benchmark, network impairments, QoS**

## I. INTRODUCTION

Surgical teleoperation has the potential to expand access to specialist surgical care. The clinical need remains substantial, with the Lancet Commission on Global Surgery estimating that around 5 billion people lack access to safe and affordable surgical care [1]. Robot-assisted telesurgery has progressed from landmark demonstrations such as the Lindbergh operation toward renewed clinical and translational interest [2], [3]. Yet practical deployment depends not only on robotic capability, but also on whether remote visual feedback remains sufficiently faithful, continuous, and timely under imperfect network conditions.

*Z. Deng, Z. Yuan, and T. Lu are with the School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom (e-mail: {zexin.deng, zhenhui.yuan, lu.tian.2}@warwick.ac.uk). Z. Yuan is a consultant member at the Institute for Applied and Translational Technologies in Surgery, University Hospitals Coventry&Warwickshire/NHS Trust, UK.

†G. Li is with the College of Control Science and Engineering and the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (e-mail: gaofeng.li@zju.edu.cn).

‡M. Huang is with Pengcheng Laboratory, Shenzhen 518000, China, and also with South China University of Technology, Guangzhou 511442, China (e-mail: huangmp@pcl.ac.cn).

§L. Zou is with Pengcheng Laboratory, Shenzhen 518000, China, and also with Southern University of Science and Technology, China (e-mail: zoulh@pcl.ac.cn).

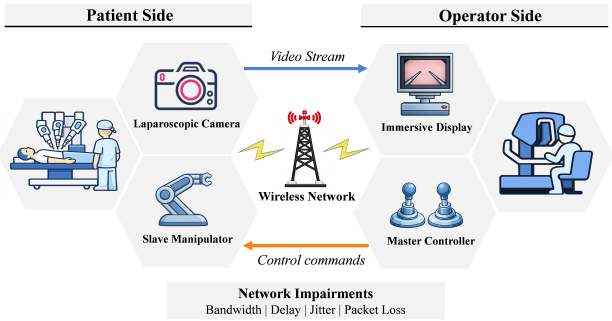



Fig. 1. Communication asymmetry in surgical teleoperation, with laparoscopic video transmitted through the forward video path and motion commands through the reverse control path.

A central communication challenge in telesurgery lies in the asymmetry between its two data paths, as illustrated in Fig. 1. For human-in-the-loop operation, system usability depends critically on the fidelity, continuity, and timeliness of the forward video stream [4]. When latency, jitter, and packet loss increase, or available bandwidth becomes constrained, the surgeon may experience spatial distortion, temporal freezing, or stale scene updates, which in turn degrade hand–eye coordination and task execution [5], [6]. In this setting, degraded video is not merely a quality-of-service (QoS) concern; it directly shapes teleoperation performance and safety.

Despite growing interest in this problem, rigorous and reproducible evaluation remains limited. Prior studies suggest that adverse communication conditions can impair telesurgical performance, but they remain difficult to compare because many vary only one or two network factors, focus primarily on QoS or task performance, or rely on software-only simulation and non-surgical settings [4], [5], [7]. As a result, there is still no reproducible benchmark that jointly provides forward-path isolation, scenario-based impairment combinations, and aligned measurement of network QoS, video quality, temporal continuity, and task performance. Table I highlights this gap, which VISTA is designed to address through a publicly released benchmark.

To address this gap, we present *VISTA* (**V**ideo **I**mpairment **S**cenario **T**estbed for Surgical Teleoperation **A**ssessment), a hardware-in-the-loop (HIL) benchmark for evaluating video feedback under network impairments in robot-assisted laparoscopic teleoperation. VISTA isolates the forward video path, keeps the reverse control path stable, and enables systematic analysis of how network impairments propagate to received video quality, temporal continuity, and task performance. The benchmark is grounded in an FLS-based peg transfer task and records aligned measurements of configured impairments,

TABLE I
COMPARISON WITH REPRESENTATIVE PRIOR WORK IN SURGICAL TELEOPERATION.

| Work | Platform | Network Impairment | Video Quality Evaluation | Task Performance | Reproducibility |
|---|---|---|---|---|---|
| Xu et al. [8] | Simulator (dV-Trainer) | Latency | Subjective | Time, Errors | Closed |
| Kumcu et al. [9] | Manual box (Non-robotic) | Latency | None | Time, Usability | Closed |
| Akasaka et al. [4] | HIL (Proprietary robot) | Bandwidth, Packet loss | Subjective | Time, Fatigue | Closed |
| Wang et al. [5] | HIL (Proprietary robot) | Bandwidth, Latency | Subjective | Time, Workload | Closed |
| **VISTA (ours)*** | **HIL (Generic with Adapter)** | **Multi-factor** | **Objective VQA** | **Time, Success Rate, Failure Analysis** | **Open (Data + CAD)** |

*VISTA employs a **universal mechanical adapter**, with the original Creo CAD files released in this work, to reduce dependence on proprietary hardware platforms. It further evaluates **multi-factor network impairments** across five realistic tiers and uses **objective VQA** (PSNR, SSIM, VMAF) instead of subjective rating scales.

measured QoS, video quality, temporal continuity, and task outcome. Our main contributions are as follows:

1) **A benchmark for video streaming performance under network impairments:** We develop a controlled network emulation platform based on Linux *tc-netem* and virtual Ethernet pairs, and define five benchmark tiers: Hospital LAN, 5G Urban, 4G Rural, LEO Satellite, and GEO Satellite, which jointly model bandwidth, latency, jitter, and bursty packet loss using a Gilbert–Elliott model. Across 375 trials of an FLS-based teleoperation task, we identify three distinct failure modes: visual degradation in 4G Rural scenarios, persistent delay-induced temporal mismatch in GEO Satellite conditions, and episodic temporal instability in LEO Satellite environments. The benchmark protocol and dataset are released to support reproducible research on network-aware teleoperation systems.
2) **An integrated evaluation pipeline:** VISTA provides a unified measurement framework that captures network QoS, objective video quality (PSNR, SSIM, VMAF), as well as temporal continuity indicators including one-way delay and freeze rate. It complements these with end-to-end task performance measurements from human-in-the-loop teleoperation trials. By aligning these heterogeneous metrics within a single evaluation pipeline, VISTA enables a consistent and reproducible analysis of how network impairments propagate through the communication stack and ultimately affect human teleoperation performance.

## II. HARDWARE-IN-THE-LOOP TESTBED AND BENCHMARK DESIGN

This section presents VISTA, which isolates the forward video path and enables reproducible evaluation of how network impairments affect video quality and task performance in surgical teleoperation.

### A. System Overview

Figure 2 illustrates the overall VISTA platform. The surgical workspace is provided by an ASPIRE Surghero laparoscopic box trainer, whose integrated endoscopic camera captures RGB video of the manipulation scene and streams it to a workstation equipped with an *NVIDIA RTX 4060* GPU for packet reception, packet assembly, decoding, visualisation, and metric logging. The decoded stream is rendered through a fixed *VIVE Focus 3* headset pipeline based on *VIVE Hub* and *SteamVR*, which remains unchanged across all trials. The operator issues motion commands through a VR controller connected to the same workstation, and these commands are forwarded over a dedicated low-latency Gigabit Ethernet link to a 6-DoF *UFactory Lite 6* robot arm. This separation between the impaired forward video path and the stable reverse control path ensures that robot actuation is unaffected by degradation on the video path. Controller pose and button state are sampled at 120 Hz through the OpenVR API.

### B. Network Impairment Injection

Controlled impairments are introduced only on the forward video path through an HIL emulation pipeline implemented on a single workstation. Video packets traverse a Linux virtual Ethernet (veth) pair and an isolated network namespace, enabling reproducible manipulation of bandwidth, delay, jitter, and packet loss without dedicated external networking hardware; the impaired stream is then received, decoded, and rendered through the fixed headset pipeline, while the reverse control path bypasses the impairment layer entirely. Network constraints are applied using Linux Traffic Control (tc) with the NetEm queue discipline [10], with delay, jitter, packet loss, and rate limiting configured atomically in a single tc qdisc rule on the veth-host interface to ensure consistency across trials. Bandwidth shaping uses NetEm's native rate parameter rather than a separate token bucket filter, avoiding cross-qdisc interaction effects. This design provides a compact, reproducible, software-defined impairment layer suitable for benchmark release.

### C. Network Tier Configuration

To emulate representative operating conditions on the forward video path, VISTA defines five benchmark network tiers: Hospital LAN, 5G Urban, 4G Rural, LEO Satellite, and GEO Satellite. Each tier is specified by joint constraints on bandwidth, one-way delay, jitter, and packet loss, so that the benchmark covers both baseline conditions with minimal impairment and progressively more challenging long-distance links.

To capture the temporally correlated burst losses characteristic of wireless links, packet loss is modeled using the simplified Gilbert model implemented in NetEm gemodel [10], [11]. In the simplified form used here, the Good state is lossless, whereas the Bad state is fully lossy. This two-state Markov model is adopted as a parsimonious approximation

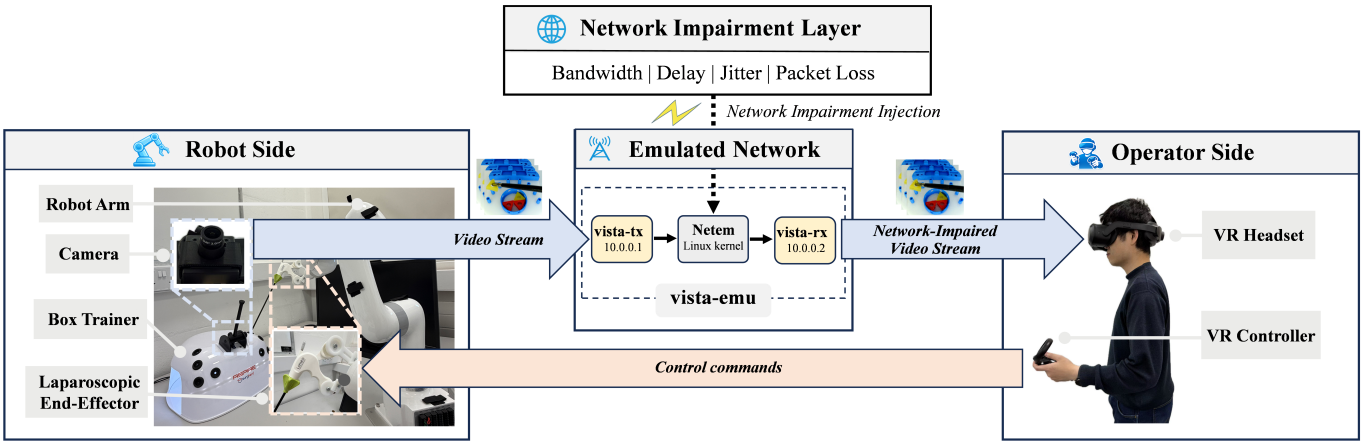



Fig. 2. Overview of the VISTA HIL testbed at the University of Warwick. The emulated network injects impairments on the forward video path between the laparoscopic box trainer and the operator side, while control commands are transmitted to the UFactory Lite 6 manipulator over a dedicated unimpaired low-latency link.

of bursty packet loss behaviour with analytically tractable parameterization: rather than attempting to reproduce the full physical-layer behaviour of wireless links, it enables direct control over two impairment characteristics that are particularly consequential for real-time surgical video, namely the steady-state loss rate and the mean burst length. Accordingly,

$$p_E = \frac{p}{p+r}, \qquad L_B = \frac{1}{r}, \tag{1}$$

where $p$ and $r$ denote the Good→Bad and Bad→Good transition probabilities, respectively [11], [12]. For each tier, the transition probabilities $p$ and $r$ are derived analytically from the target steady-state loss rate $p_E$ and the target mean burst length $L_B$ for that tier.

Table II summarises the benchmark network tiers. The parameter values were chosen to span representative teleoperation conditions, from near ideal local connectivity to severely constrained long-distance links, and were grounded in standard performance objectives, published measurements, and clinical latency constraints. Hospital LAN serves as a strict dedicated fibre baseline and lies well within ITU-T Y.1541 objectives [13]. 5G Urban is included as a conservative 5G tier informed by the service requirements in 3GPP TS 22.261 and recent telesurgery studies [3], [14]. 4G Rural represents a degraded cellular tier near the operating margin identified in [5]. LEO Satellite models a Starlink class low Earth orbit link, with periodic delay and burst loss excursions added to the static parameters in the table, consistent with reported Starlink behaviour [15]. GEO Satellite, with a one-way delay of 600 ms, is included as a high latency reference condition that lies well beyond the guidance for highly interactive applications and exceeds the 320 ms clinical ceiling reported in [5], [16].

## III. EXPERIMENTAL PROTOCOL AND METRICS

This section describes the task protocol, video transport settings, and aligned measurement pipeline used in VISTA. Each trial yields a single record spanning configured impairments, measured QoS, received video quality, temporal continuity, and task outcome.

(a) Workspace definition (b) Initial state (c) Target placement

Fig. 3. Central-to-Peripheral Peg Transfer (C2P) task used in VISTA. Four perforated triangular objects are transferred from the central loading zone $C_0$ to the target pegs $P_1$–$P_4$ under network-impaired video feedback.

### A. Task Definition and Operator Preparation

We adopt a Central-to-Peripheral Peg Transfer (C2P) task derived from the Fundamentals of Laparoscopic Surgery (FLS) peg transfer task. Unlike the canonical bimanual FLS task, the protocol is performed with a teleoperated single-arm *UFactory Lite 6* manipulator under a fixed laparoscopic camera view, preserving the core demands of visually guided grasping, transfer, alignment, and release.

As shown in Fig. 3, the central circular region $C_0$ serves as the loading zone and four peripheral pegs $P_1$–$P_4$ serve as target locations. At the start of each trial, four perforated triangular objects are placed in $C_0$. The operator grasps each object under network-impaired video feedback, transfers it within the same field of view, aligns its hole with the designated peg, and releases it. A trial ends when all four objects are placed or when the predefined time limit of 300 s is reached.

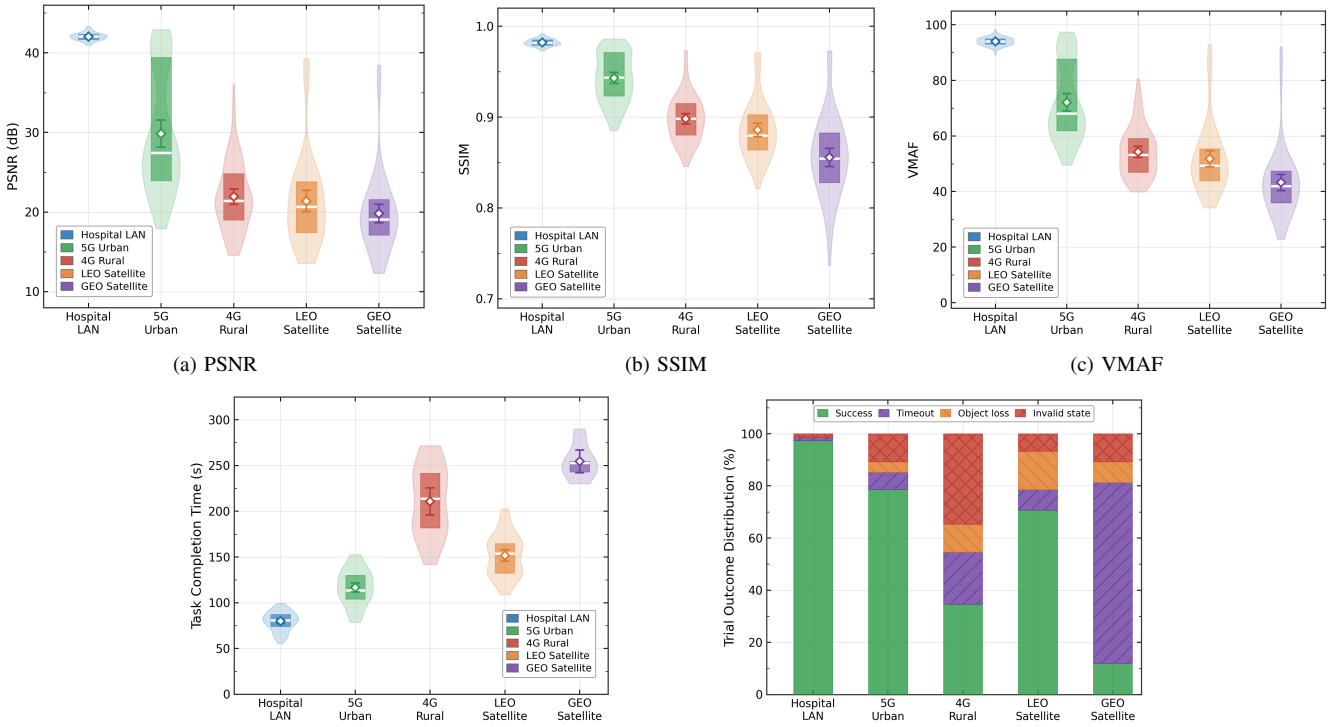
To reduce learning effects, participants completed standardised pre-experimental training in the Marion Surgical Robot Simulator.[1] The simulator was used as a proficiency

[1]Marion Surgical Robot Simulator: https://store.steampowered.com/app/3457580/Marion_Surgical_Robot_Simulator/

TABLE II
BENCHMARK NETWORK TIERS AND IMPAIRMENT PARAMETERS.

| Networks | Bandwidth (Mbps) | One-way delay (ms) | Jitter (ms) | Packet loss | | | |
|---|---|---|---|---|---|---|---|
| | | | | $p$ (%) | $r$ (%) | $p_E$ (%) | $L_B$ (pkts) |
| Hospital LAN | 100 | 3 | 1 | 0.005 | 50.0 | 0.01 | 2 |
| 5G Urban | 20 | 35 | 8 | 0.167 | 33.3 | 0.50 | 3 |
| 4G Rural | 5 | 150 | 30 | 0.341 | 16.7 | 2.00 | 6 |
| LEO Satellite | 30 | 30 | 10 | 0.305 | 20.0 | 1.50 | 5 |
| GEO Satellite | 10 | 600 | 50 | 0.251 | 50.0 | 0.50 | 2 |

*Note:* $p$ and $r$ denote the Good→Bad and Bad→Good transition probabilities passed to tc netem gemodel, respectively; $p_E$ is the steady-state loss rate, and $L_B$ is the mean burst length.




(a) PSNR (b) SSIM (c) VMAF

(d) Task Completion Time (success only) (e) Trial Outcome Distribution

Fig. 4. Results across the five benchmark tiers. Panels (a)–(c) show PSNR, SSIM, and VMAF; panel (d) shows task completion time for successful trials; and panel (e) shows the distribution of trial outcomes. In panels (a)–(d), the coloured box denotes the interquartile range, the white line denotes the median, and the diamond denotes the mean ± 95 % CI. Numerical summary is given in Table III.

filter rather than a surrogate for surgical competence [17]. Training metrics included Grasper Collisions, Distance Travelled, and Offscreen Time. Operators qualified for the HIL study only after meeting the predefined criterion in three consecutive sessions: zero grasper collisions and inter-session variation in travelled distance below $\pm 5\%$ [18]. The study was approved by the University of Warwick Biomedical and Scientific Research Ethics Committee (BSREC 64/25-26), and all participants provided informed consent.

A total of 15 qualified operators completed 375 trials. Each operator performed 5 trials per condition across all five tiers under fixed codec settings, yielding 75 trials per tier. Condition order was counterbalanced across operators, with short rest periods between blocks.

### B. Video Transport and Receiver Configuration

The video source is the integrated endoscopic camera of the ASPIRE Surghero trainer, producing $640 \times 480$ RGB video at 30 fps over USB/V4L2. H.264 (libx264, GOP = 15 frames, tune=zerolatency) is used as the sole codec across all tiers, reflecting predictive coding in practical teleoperation video systems [5]. The tune=zerolatency configuration produces one access unit per frame, while the fixed GOP limits error propagation to at most 15 frames. Each encoded access unit is fragmented into 8 000-byte UDP chunks with a 16-byte custom header containing the frame ID, chunk index, total chunk count, and a 64-bit capture timestamp. Because sender and receiver run on the same workstation, clock synchronization uncertainty is removed from one-way timestamping, and an adaptive reassembly buffer is used at the receiver to mitigate false frame drops under high jitter. On H.264 decode failure, the most recent valid frame is redisplayed and the event is logged as a freeze; concealed frames are excluded from PSNR and SSIM calculations. Trial timing starts on receipt of the first decoded frame, yielding an alignment offset bounded by

TABLE III
KEY RESULTS ACROSS THE FIVE BENCHMARK TIERS.

| Networks | Completion Time (s)* | Success Rate (%) | Successful Trials | PSNR (dB) | SSIM | VMAF | Freeze Rate (events/min) |
|---|---|---|---|---|---|---|---|
| Hospital LAN | $80 \pm 2$ | 97 | 73 | 42.0 | 0.982 | 94 | 0.0 |
| 5G Urban | $117 \pm 5$ | 79 | 59 | 29.9 | 0.943 | 72 | 1.1 |
| 4G Rural | $211 \pm 15$ | 35 | 26 | 21.9 | 0.898 | 54 | 4.5 |
| LEO Satellite | $152 \pm 6$ | 71 | 53 | 21.4 | 0.886 | 52 | 3.4 |
| GEO Satellite | $255 \pm 12$ | 12 | 9 | 19.8 | 0.856 | 43 | 5.1 |

*Completion time is reported as mean $\pm$ 95% CI for successful trials only.

one frame period. Keeping codec settings constant across all tiers ensures that quality variation is attributable to network impairment rather than encoder adaptation. For objective video quality assessment, the source and received videos were automatically recorded in each trial, and PSNR, SSIM, and VMAF were computed offline using the Video Quality Measurement Tool.[2]

## IV. RESULTS AND ANALYSIS

Trial-level metrics were aggregated across the 75 trials in each tier, and means are reported with 95% confidence intervals (CIs). Within each trial, PSNR, SSIM, and VMAF were averaged over successfully decoded frames. Freeze rate, success rate, failure mode, and completion time were recorded at the trial level.

### A. Overall Performance

Fig. 4 and Table III show a clear progression from near-ideal operation to severe teleoperation impairment. Hospital LAN provides the reference baseline; 5G Urban remains largely usable, although its bimodal quality distribution suggests intermittent H.264 degradation.

The three more impaired tiers present a more important result: similar average frame quality does not imply similar task performance. 4G Rural, LEO Satellite, and GEO Satellite reach mean PSNR values of 21.9, 21.4, and 19.8 dB, respectively, yet their success rates diverge substantially to 35%, 71%, and 12%, and their mean completion times for successful trials span from $152 \pm 6$ s to $255 \pm 12$ s. This is consistent with prior telesurgery studies showing that communication constraints affect performance through more than one pathway, including both latency and video degradation [4], [5].

### B. Network Impairments and Task Performance

The results reveal that similar mean frame quality does not produce similar task outcomes, and that the dominant impairment mechanism differs across tiers.

In 4G Rural, the primary driver is sustained spatial quality degradation: low bandwidth (5 Mbps) combined with bursty packet loss ($p_E$=2%, $L_B$=6 pkts) reduces mean PSNR to 21.9 dB and raises the freeze rate to 4.5 events/min. Under these conditions, operators frequently lose visual contact with the instrument tip, leading to a success rate of only 35% and the highest proportion of failures due to invalid state among all tiers. The mean completion time for the minority of successful trials reaches $211 \pm 15$ s, reflecting the additional time required to recover from degraded spatial feedback.

In GEO Satellite, the dominant mechanism is fundamentally different: despite a mean PSNR of 19.8 dB comparable to 4G Rural, the 600 ms one-way delay means that visual feedback can lag the physical scene by more than one second over a full perception–action cycle. This persistent temporal mismatch between the displayed and actual instrument position causes the operator to act on stale imagery, producing the lowest success rate across all tiers (12%) and the longest mean completion time ($255 \pm 12$ s). Timeout is the dominant outcome (69%), consistent with operators repeatedly correcting positional errors induced by outdated visual feedback rather than failing to perceive the instrument entirely. The one-way delay of 600 ms substantially exceeds both the 320 ms clinical ceiling reported in [5] and the broader latency limits discussed in telesurgery studies [6], [8].

In LEO Satellite, mean PSNR (21.4 dB) and overall success rate (71%) suggest moderate impairment, yet the elevated standard deviation of completion time and the 15% object-loss rate indicate that degradation is episodic rather than sustained. Analysis of individual LEO trials reveals distinct impairment episodes coinciding with frame loss bursts and PSNR reduction. The most severe episodes reach concealed-frame proportions exceeding 85%, before recovering to baseline within seconds. These transient episodes disrupt continuous hand–eye coordination precisely when fine alignment is required, leading to object loss even when average conditions remain less severe than in GEO Satellite or 4G Rural.

Taken together, these findings show that network impairments degrade task performance through at least two distinct pathways: sustained spatial quality loss due to bandwidth limitation and packet loss (dominant in 4G Rural), and temporal mismatch between displayed and actual instrument state caused by high propagation delay (dominant in GEO Satellite), with episodic instability (LEO Satellite) constituting a third, qualitatively different mode. Mean frame quality metrics alone cannot distinguish these pathways, which motivates the multi-metric evaluation pipeline provided by VISTA.

### C. Failure Analysis

The outcome distributions in Fig. 4e reveal that the three impaired tiers fail in qualitatively different ways. In 4G Rural, the dominant failure is invalid state (35%), followed by timeout (20%): low spatial quality and a freeze rate of 4.5 events/min

[2]Video Quality Measurement Tool (VQMT): https://videoprocessing.ai/vqmt/

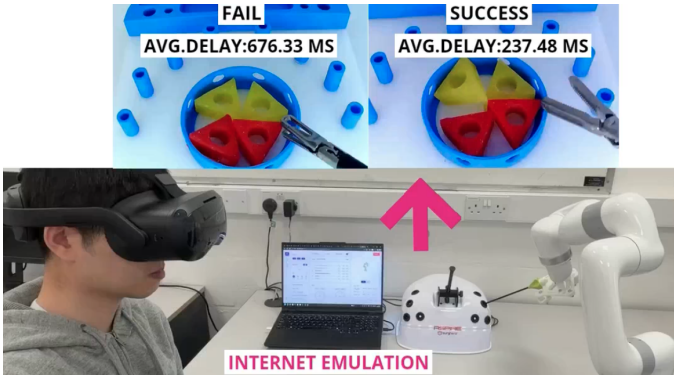



Fig. 5. Representative trial outcomes under GEO Satellite (left, failed, average OWD 676 ms) and 4G Rural (right, successful, average OWD 237 ms), illustrating that high propagation delay and spatial quality loss constitute fundamentally different failure mechanisms.

directly impair the fine motor control required for peg alignment. In GEO Satellite, timeout dominates (69%), as operators repeatedly issue corrective commands based on stale imagery, creating a compounding positional error cycle that exhausts the 300 s limit. In LEO Satellite, success remains 71%, but object loss rises to 15%, reflecting episodic disruption to hand–eye coordination rather than sustained degradation.

Fig. 5 contrasts a representative failed GEO Satellite trial (average one-way delay 676 ms) with a successful 4G Rural trial (237 ms): despite comparable mean PSNR, the persistent temporal mismatch in GEO Satellite prevented task completion, while the 4G Rural operator recovered from spatial degradation and succeeded.

Taken together, 4G Rural and LEO Satellite reach comparable mean PSNR yet differ by 36 percentage points in success rate, while GEO Satellite and 4G Rural share similarly low mean PSNR yet exhibit entirely different dominant failure modes. These results confirm that spatial fidelity, propagation delay, and the temporal pattern of impairment must be considered jointly for teleoperation safety evaluation.

## V. Conclusion

In this paper, we presented VISTA, a hardware-in-the-loop benchmark for evaluating video feedback under network impairments in robot-assisted laparoscopic teleoperation. By isolating the forward video path and emulating bandwidth, latency, jitter, and burst loss in a controlled manner, VISTA enables reproducible evaluation of received video quality, temporal continuity, and task performance. Results from 375 trials across five representative network tiers showed that similar average frame quality does not necessarily lead to similar teleoperation outcomes. VISTA therefore provides a practical basis for evaluating how spatial quality, one-way delay, and the temporal pattern of network impairment jointly affect teleoperation performance.

## References

[1] J. G. Meara, A. J. M. Leather, L. Hagander, B. C. Alkire, N. Alonso, E. A. Ameh, S. W. Bickler, L. Conteh, A. J. Dare, J. Davies, *et al.*, "Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development," *Lancet*, vol. 386, no. 9993, pp. 569–624, 2015.

[2] J. Marescaux, J. Leroy, M. Gagner, F. Rubino, D. Mutter, M. Vix, S. E. Butner, and M. K. Smith, "Transatlantic robot-assisted telesurgery," *Nature*, vol. 413, no. 6854, pp. 379–380, 2001.

[3] B. Rocco, M. C. Moschovas, S. Saikali, G. Gaia, V. Patel, and M. C. Sighinolfi, "Insights from telesurgery expert conference on recent clinical experience and current status of remote surgery," *Journal of Robotic Surgery*, vol. 18, no. 1, p. 240, 2024.

[4] H. Akasaka, K. Hakamada, H. Morohashi, T. Kanno, K. Kawashima, Y. Ebihara, E. Oki, S. Hirano, and M. Mori, "Impact of the suboptimal communication network environment on telerobotic surgery performance and surgeon fatigue," *PLoS One*, vol. 17, no. 6, p. e0270039, 2022.

[5] Y. Wang, Q. Ai, T. Shi, B. Gao, W. Zhao, C. Jiang, G. Liu, L. Zhang, H. Li, et al., "Influence of network latency and bandwidth on robot-assisted laparoscopic telesurgery: A pre-clinical experiment," *Chinese Medical Journal*, vol. 138, no. 3, pp. 325–331, 2025.

[6] Z. Y. Motiwala, A. Desai, R. Bisht, S. Lathkar, S. Misra, and D. D. Carbin, "Telesurgery: current status and strategies for latency reduction," *J. Robot. Surg.*, vol. 19, no. 1, p. 153, 2025.

[7] Z. Deng, Z. Yuan, and L. Zou, "TeleSim: A Network-Aware Testbed and Benchmark Dataset for Telerobotic Applications," *arXiv preprint arXiv:2507.04425*, 2025.

[8] S. Xu, M. Perez, K. Yang, C. Perrenot, J. Felblinger, and J. Hubert, "Determination of the latency effects on surgical performance and the acceptable latency levels in telesurgery using the dV-Trainer® simulator," *Surg. Endosc.*, vol. 28, no. 9, pp. 2569–2576, Sep. 2014, doi: 10.1007/s00464-014-3504-z.

[9] A. Kumcu, L. Vermeulen, S. A. Elprama, P. Duysburgh, L. Platiša, Y. Van Nieuwenhove, N. Van De Winkel, A. Jacobs, J. Van Looy, and W. Philips, "Effect of video lag on laparoscopic surgery: correlation between performance and usability at low latencies," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 13, no. 2, Art. no. e1758, 2017, doi: 10.1002/rcs.1758.

[10] F. Ludovici and H. P. Pfeifer, "tc-netem(8) — Linux manual page," *man7.org*. [Online]. Available: https://www.man7.org/linux/man-pages/man8/tc-netem.8.html. Accessed: Mar. 14, 2026.

[11] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, no. 5, pp. 1253–1265, 1960.

[12] G. Haßlinger and O. Hohlfeld, "The Gilbert–Elliott model for packet loss in real-time services on the Internet," in *Proc. 14th GI/ITG Conf. Meas., Model. Eval. Comput. Commun. Syst. (MMB)*, Dortmund, Germany, 2008, pp. 1–15.

[13] ITU-T, "Network performance objectives for IP-based services," *ITU-T Rec. Y.1541*, 2011.

[14] 3GPP, "Service requirements for the 5G system," *3GPP TS 22.261*, Release 17.

[15] R. Richter, V. Ververis, and V. Bajpai, "Breaking Through the Clouds: Performance Insights into Starlink's Latency and Packet Loss," in *IFIP Networking Conference (Networking)*, 2025.

[16] ITU-T, "One-way transmission time," *ITU-T Rec. G.114*, 2003.

[17] A. Moglia, V. Ferrari, L. Morelli, M. Ferrari, F. Mosca, and A. Cuschieri, "A systematic review of virtual reality simulators for robot-assisted surgery," *Eur. Urol.*, vol. 69, no. 6, pp. 1065–1080, 2016.

[18] N. J. Hogle, W. D. Widmann, A. O. Ude, M. A. Hardy, and D. L. Fowler, "Does training novices to criteria and does rapid acquisition of skills on laparoscopic simulators have predictive validity or are we just playing video games?," *J. Surg. Educ.*, vol. 65, no. 6, pp. 431–435, 2008.