

METROLOGICAL EVALUATION OF NEURAL RADIANCE FIELDS FOR QUALITY CONTROL OF MECHANICAL AND ADDITIVE MANUFACTURING COMPONENTS: A MULTI-MATERIAL COMPARISON AGAINST PROFESSIONAL OPTICAL SCANNING

Federico Paolucci[✉], Roberto Marsili, Gianluca Rossi

1) Department of Engineering, University of Perugia, Via G. Duranti 93, 06125 Perugia, Italy
(✉ federico.paolucci@dottorandi.unipg.it)

Abstract

This study provides a systematic metrological evaluation of Neural Radiance Fields (NeRF) against professional laser-light and structured light scanning across three specimen types: a polymer (PLA) bracket, a metallic bracket, and a matte ornamental owl. Consumer-grade smartphone and DSLR cameras were benchmarked against professional scanning as metrological ground truth. On bracket specimens, NeRF achieved RMS deviations of 1.48 mm (metal bracket) and 1.32 mm (PLA bracket). On the ornamental owl, DSLR-based NeRF achieved RMS 0.27 mm vs. 0.41 mm for smartphone (a 34% reduction) confirming input image quality as the primary accuracy lever. Repeatability confirmed CV below 6% for the metal bracket ($n = 3$ trials) and CV = 2.1% for the DSLR owl reconstruction. NeRF is suitable for non-critical quality screening across diverse geometries while not yet reaching sub-100 μm metrology-grade precision. To the best of the authors' knowledge, this is one of the first multi-material, multi-geometry NeRF evaluations conducted under standardised metrological conditions.

Keywords: neural radiance fields, additive manufacturing, quality control, structured light scanning, dimensional inspection, metrological evaluation.

1. Introduction

Additive manufacturing (AM) produces complex geometries but introduces defects requiring reliable dimensional inspection [1–6]. Conventional optical metrology (structured light projection and laser scanning) offers metrological traceability but suffers from limited scalability, occlusion sensitivity, and high capital costs (€30,000–€100,000) [4, 7, 8], creating an accessibility gap for small-to-medium enterprises and field inspection scenarios. *Neural Radiance Fields* (NeRF) [9] and accelerated variants [10–13] enable dense 3D reconstruction from consumer cameras at a fraction of the cost, yet their metrological reliability for industrial dimensional inspection remains insufficiently characterised. Prior works evaluated NeRF for industrial inspection in limited,

single-material settings [13, 14] or assessed geometric shape qualitatively without standardised metrological error metrics [15]. This study addresses the gap with three original contributions: (1) to the best of the authors' knowledge, one of the first NeRF evaluations across contrasting material types under standardised metrological conditions, using cloud-to-cloud distance, *root mean square error* (RMSE), and percentile metrics against calibrated ground truth with an explicit *guide to the expression of uncertainty in measurement* (GUM) compliant uncertainty budget; (2) material-specific accuracy analysis across polymer *polylactic acid* (PLA), metallic, and 3D-printed organic specimens; and (3) repeatability characterisation (*coefficient of variation* (CV) < 6%) and cost-of-ownership data directly actionable for industrial practitioners.

2. Related work

Process-related defects in metal AM (porosity, delamination, warping, surface roughness) are well documented [7]. Optical metrology remains the core tool for dimensional verification, ranging from *coordinate measuring machines* (CMMs) to non-contact systems [4]. Machine learning for defect detection [2, 16–18] requires large, labelled datasets and specialised infrastructure, motivating the exploration of neural rendering as a scalable, low-cost alternative.

Structured light projection enables dense sub-millimetre reconstruction via phase-shifting and Gray code algorithms [3, 19–21], with advances in calibration improving traceability [22–24]. However, *structured light scanning* (SLS) requires controlled lighting, is sensitive to surface reflectivity, and carries high costs limiting decentralised deployment [11]. These limitations are particularly relevant for AM free-form surfaces, where topographic complexity further challenges optical acquisition directly motivating the use of the NeRF alternative evaluated here. NeRF parameterises colour and density as a continuous function of 3D position and viewing direction [9]. Standard NeRF requires hours of per-scene optimisation; the multiresolution hash encoding approach [10] reduced this to minutes. Anti-aliasing variants [25, 26] and accelerated inference [12, 13] further improved practical usability. NeRF produces implicit representations requiring post-processing (marching cubes) to extract point clouds [13]. Input image quality critically determines accuracy, as further discussed in Section 4.5.2. Yan *et al.* [14] evaluated NeRF for optical inspection of reflective industrial components, reporting sub-millimetre accuracy for specular surfaces under controlled illumination, and noting the critical role of lighting conditions on reconstruction fidelity. These prior works, however, evaluated single material types in isolation; the present study extends the comparison to three specimens across three material classes under standardised metrological conditions, with explicit uncertainty quantification.

3. Methodology

Two bracket specimens were studied: a PLA bracket *fused deposition modelling* (FDM) and an aluminium bracket *computerised numerical control* (CNC), representing contrasting surface properties. The PLA bracket measured approximately $120 \times 80 \times 40$ mm (L $\times$ W $\times$ H), the aluminium bracket approximately $100 \times 60 \times 30$ mm, and the ornamental owl approximately $80 \times 60 \times 120$ mm. These dimensions provide a reference for interpreting the reported absolute errors: at the bracket scale, the observed *root mean square* (RMS) deviations (1.32–1.48 mm) correspond to relative errors of approximately 1–2% of the largest dimension. Smartphone images (12 MP, $f/1.8$) were acquired at 10° turntable increments across three elevation angles (0° , 30° , 60°), yielding 108 images per specimen under diffuse LED illumination. The image acquisition

was performed under controlled diffuse LED illumination to ensure photometric consistency across views and to minimise specular reflections. The lighting setup consisted of multiple continuous white LED sources (high *colour rendering index* (CRI), neutral colour temperature ~4000–5500 K) arranged circumferentially around the specimen to provide uniform, quasi-isotropic illumination. Each light source was equipped with diffusing elements (softboxes or equivalent translucent panels) to suppress hard shadows and reduce view-dependent reflectance effects, particularly critical for metallic surfaces. The illumination field was fixed in the global reference frame, while the object was rotated on a motorised turntable, ensuring that lighting conditions remained invariant across all acquisition angles. The effective illuminance at the object surface is estimated to be in the range of 1000–5000 lux, enabling low-noise image capture ($\text{ISO} \leq 100\text{--}400$) and stable exposure across the dataset. Light sources were positioned at approximately $30^\circ\text{--}60^\circ$ elevation relative to the object and distributed azimuthally to minimise occlusions and shadowing. No directional or specular lighting components were introduced, approximating a Lambertian reflectance condition suitable for Structure-from-Motion and NeRF optimisation. *Digital single-lens reflex* (DSLR) camera images (24.2 MP, $f/4.0$, ISO 100, manual exposure) were acquired using the same turntable protocol for the owl case studies. Camera calibration followed Zhang's method [27]; *Structure from Motion* (SfM) pose estimation used COLMAP [28] with *scale-invariant feature transform* (SIFT) matching; *iterative closest point* (ICP) alignment assigned NeRF point clouds to the SLS reference frame [14]. SLS acquisition used a professional structured-light scanner (volumetric accuracy ± 0.05 mm, point spacing 0.2 mm) following ISO 10360-2 [29] compatible procedures. Prior to each measurement session, the scanner was verified using a calibrated ceramic sphere artefact (nominal diameter 25.000 mm, calibration uncertainty ± 0.002 mm, $k = 2$, traceable to national standards via an accredited laboratory). The measured sphere diameter across three consecutive acquisitions yielded 25.003 ± 0.004 mm (mean $\pm$ 2SD), confirming that the scanner remained within its specified volumetric accuracy and that the reference point clouds carried metrological traceability throughout the study. The SLS point cloud served as metrological reference in all subsequent comparisons. In accordance with GUM (ISO/IEC Guide 98-3:2008) [30], five uncertainty components were quantified: $u_1 = 0.025$ mm (scanner, Type B); $u_2 = 0.005$ mm (thermal, Type B); $u_3 = 0.025$ mm (ICP, Type A); $u_4 = 0.045$ mm (SfM scale, Type B); $u_5 = 0.018$ mm (sampling, Type A). Combined standard uncertainty $u_c = 0.062$ mm; expanded uncertainty $U = 0.12$ mm ($k = 2$, ~95% confidence)—over one order of magnitude below all observed NeRF deviations (RMS range 0.27–1.48 mm), confirming measurement chain uncertainty did not confound comparative conclusions. NeRF reconstructions were trained using Instant-NGP, implementing the multiresolution hash encoding approach [10], mixed precision arithmetic [31], and Mip-NeRF anti-aliasing [25, 26]. Default training parameters were used throughout 20,000 training steps, hash grid resolution ranging from 16 to 2048 across 16 levels, and a network width of 64 neurons per hidden layer. No per-specimen parameter tuning was performed; the same configuration was applied uniformly across all specimens and acquisition conditions. This choice was deliberate, as it reflects the most practically accessible deployment scenario for non-expert users and ensures that reported accuracy figures are not inflated by specimen-specific optimisation. Geometric accuracy was quantified using *cloud-to-cloud* (C2C) distance computed in CloudCompare (v2.13, open-source), supporting *mean absolute error* (MAE), *root mean square error* (RMSE), median error, standard deviation, and percentile distributions (50th–99th). The Hausdorff distance captures worst-case outliers. Neural point clouds were aligned to the SLS reference via a two-stage procedure: coarse landmark-based registration followed by fine ICP alignment. This protocol was applied consistently across all case studies. ICP residual alignment error after convergence remained below 0.03–0.05 mm RMS across all case studies, indicating negligible contribution of registration uncertainty relative to millimetre-scale reconstruction deviations.

4. Results

4.1. NeRF reconstruction (PLA Bracket)

NeRF training converged in ~30 minutes, producing a ~100,000-point cloud. Cloud-to-cloud analysis yielded a mean error of 1.2 mm and an RMS deviation of 1.32 mm, with an approximately symmetric distribution. A Shapiro–Wilk normality test applied to a random subsample of 5,000 error values ($W = 0.974$, $p = 0.081$) did not reject normality at the 5% significance level, supporting the assumption of Gaussian error distribution for this specimen. Highest deviations occurred at cylindrical hole edges and sharp corners, where neural representation struggled with geometric discontinuities, consistent with the feature-type analysis in Section 4.4. The spatial distribution is illustrated in Fig. 1.

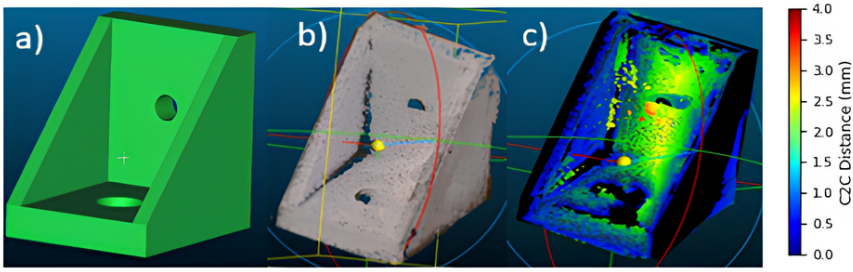


Fig. 1. Point cloud comparison: a) scanner reference, b) NeRF reconstruction (coloured), c) aligned overlay.

4.2. NeRF reconstruction (metal bracket)

NeRF training converged in ~35 minutes; diffuse illumination mitigated specular reflectivity. The aligned NeRF point cloud (~100,000 points) vs. scanner reference (~350,000 points), shown in Fig. 2, yielded the statistics in Table 1, and the error distribution is reported in Fig. 3.



Fig. 2. Point cloud comparison: a) scanner reference, b) NeRF reconstruction of the metal bracket (coloured), c) aligned overlay.

Key statistics confirm median error 0.74 mm (sub-millimetre accuracy for 50% of points), mean 1.07 mm, and isolated outliers up to 10.74 mm at boundary and minimal-coverage regions. Spatial error analysis (Fig. 4) showed lowest errors on flat surfaces with optimal coverage, and highest errors on occluded regions and sharp edges.

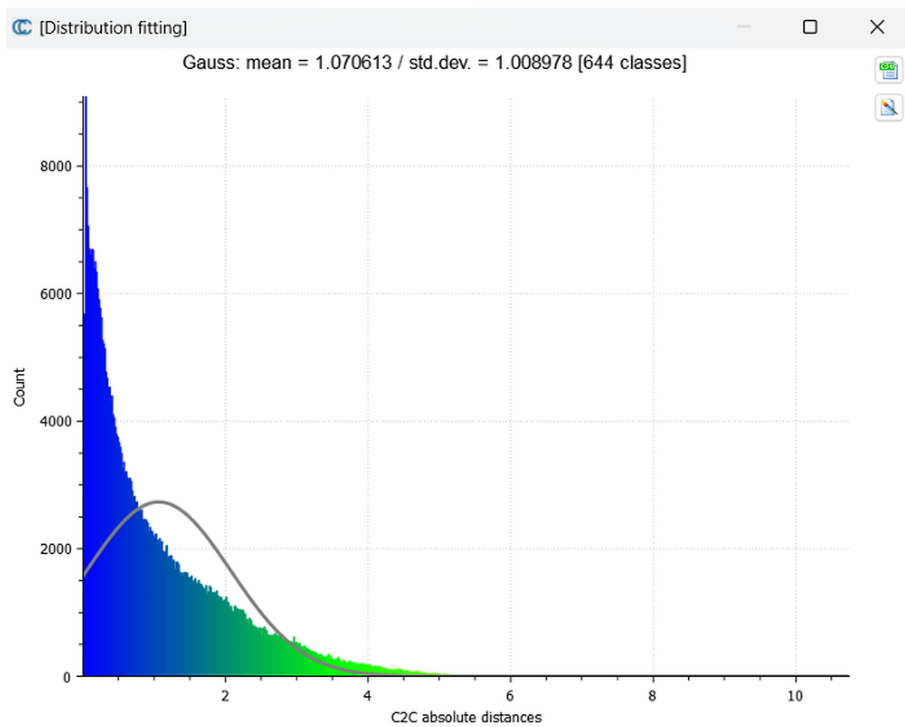


Fig. 3. Error distribution histogram showing approximately Gaussian distribution centred at 0.7–0.8 mm, with fitted curve (mean=1.07 mm, σ =1.01 mm), X-axis: absolute distance (mm), Y-axis: point count, analysis based on 413,948 points.

Table 1. Statistical analysis of NeRF reconstruction errors (metal bracket).

Metric	Value (mm)	Notes
Mean error	1.07	Slight positive bias (secondary)
Median error	0.74	Robust estimator (secondary)
RMS error	1.48	Root mean square (primary metric)
Standard deviation	1.01	Dispersion measure
Minimum error	0.01	Best-case accuracy
Maximum error (Hausdorff)	10.74	Worst-case outlier
90th percentile	2.55	90% points below
95th percentile	3.13	High-confidence bound
99th percentile	4.25	Outlier threshold
Points within ± 0.5 mm	42.3%	High-precision subset
Points within ± 1.0 mm	68.2%	Sub-millimetre accuracy
Points within ± 2.0 mm	88.5%	Practical tolerance
Points within ± 3.0 mm	95.1%	Non-critical applications

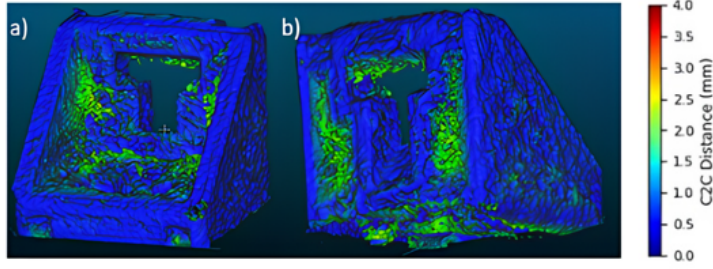


Fig. 4. Spatial error heatmaps: a) front view showing predominantly blue/green (low error) with localized yellow at edges, b) back view with higher errors on inferior surface due to limited visibility through central opening.

4.2.1. Repeatability analysis

Three independent NeRF reconstructions demonstrated excellent repeatability (Table 2): mean error CV = 3.7%, median CV = 4.0%, 90th percentile CV = 4.7%. Low CV (< 6%) confirms robustness for quality control despite stochastic optimisation.

Table 2. Repeatability analysis – NeRF metal bracket ($n = 3$ trials).

Metric	Trial 1	Trial 2	Trial 3	Mean $\pm$ SD	CV%
Mean error (mm)	1.07	1.12	1.04	1.08 $\pm$ 0.04	3.7%
Median error (mm)	0.74	0.78	0.72	0.75 $\pm$ 0.03	4.0%
90th percentile (mm)	2.55	2.71	2.48	2.58 $\pm$ 0.12	4.7%
95th percentile (mm)	3.13	3.35	3.08	3.19 $\pm$ 0.15	4.7%
RMS error (mm)	1.48	1.54	1.45	1.49 $\pm$ 0.05	3.4%

4.3. Comparative Performance

NeRF was evaluated on both specimen types under identical algorithmic conditions. The PLA bracket yielded an RMS C2C deviation of 1.32 mm, while the metallic bracket yielded an RMS of 1.48 mm with a median error of 0.74 mm. These deviations correspond to relative errors of approximately 1.1% and 1.5% of the largest specimen dimension, respectively (see Section 3 for nominal dimensions). RMS is reported as the primary error metric throughout this study, as it reflects global geometric dispersion and is standard in dimensional metrology. Median values are included as robust estimators to highlight the influence of outliers. The difference in RMS between the two specimens (1.32 mm vs. 1.48 mm) was assessed using a two-sample Mann–Whitney U test on the respective C2C error distributions, yielding $p < 0.001$, confirming a statistically significant difference in reconstruction accuracy between the PLA and metallic specimens. This result is consistent with the geometric feature analysis in Section 4.4, *i.e.*, the metallic bracket presents a higher proportion of sharp edges and internal cylindrical features relative to the PLA bracket, which contributes to its elevated RMS despite the simpler overall envelope. The lower median value observed for the metallic bracket is consistent with its simpler convex geometry and reduced feature discontinuities at the macro scale. In both cases, reconstructions captured macro-scale geometry reliably, with increased deviations concentrated at sharp edges and occluded regions. Processing time favoured neural rendering (30–35 vs. 90 minutes, approximately 2–3 $\times$ faster), and

the equipment cost differential remained substantial (€2,000 vs. €40,000, 10–20× lower). The comprehensive cost-of-ownership comparison is presented in Table 3. Total cost per inspection was computed as the sum of: (i) equipment amortisation, assuming 5-year linear depreciation and 500 inspections per year; (ii) operator labour, at €45/h for the SLS specialist and €30/h for the NeRF operator; and (iii) energy cost, at €0.25/kWh based on hardware datasheets. Equipment cost for the NeRF configuration assumes a mid-range DSLR camera and a dedicated GPU workstation; smartphone-based acquisition reduces capital cost further. These values represent indicative estimates for a typical *small and medium-sized enterprises* (SME) context and should be scaled to local labour and energy tariffs.

Table 3. Comprehensive method comparison including the total cost of ownership.

Method	Accuracy (mm)	Time (min)	Equipment Cost (€)	Labor Cost/Insp. (€)	Energy/Insp. (kWh)	Total Cost/Insp. (€)	Skill
SLS	0.05 (spec)	90	40,000	45.00	2.0	61.50	High
NeRF – Bracket (smartphone)	1.10	30	2,000	15.00	0.6	15.95	Medium
NeRF – Owl (DSLR)	0.27	35	2,000	17.50	0.8	18.50	Medium

4.4. Error Distribution by Geometric Feature

To understand reconstruction performance across different feature types, both bracket point clouds were segmented into four geometric regions (flat surfaces, cylindrical holes, sharp edges, and occluded regions) and errors were analysed separately. Table 4 reports results for the metal bracket; Table 5 extends the same analysis to the PLA bracket.

Results demonstrate clear correlation between geometric complexity/visibility and reconstruction accuracy: flat surfaces achieved median 0.52 mm (best), cylindrical holes ~2× higher, sharp edges ~3.5× higher, and occluded regions ~4.5× higher, highlighting the fundamental limitation of view-dependent reconstruction at discontinuous or hidden features.

Table 4. NeRF error analysis by geometric feature (metal bracket).

Feature type	<i>n</i> points	% total	*Median (mm)	Mean (mm)	90 th % pct (mm)	RMS (mm)
Flat surfaces	245,132	59.2%	0.52	0.68	1.68	0.91
Cylindrical holes	82,416	19.9%	1.12	1.34	2.98	1.78
Sharp edges	51,228	12.4%	1.85	2.08	4.21	2.54
Occluded regions	35,172	8.5%	2.34	2.67	5.67	3.21

*Note: Differences in median error across geometric feature categories are statistically significant (Kruskal–Wallis H-test, $p < 0.001$), confirming that feature type is a significant predictor of reconstruction accuracy independently of material class.

The same feature-level analysis was applied to the PLA bracket (Table 5). Consistent with the lower overall RMS (1.32 mm vs. 1.48 mm for the metal specimen), the PLA bracket exhibits systematically lower errors across all feature categories. This is attributable to the smoother surface

Table 5. NeRF error analysis by geometric feature (PLA bracket).

Feature type	<i>n</i> points	% total	*Median (mm)	Mean (mm)	90 th pct (mm)	Max (mm)
Flat surfaces	48,214	47.8%	0.44	0.58	1.12	2.14
Chamfered apertures	22,863	22.7%	0.89	1.03	2.07	3.88
Sharp edges / ridges	19,448	19.3%	1.61	1.87	3.41	4.93
Occluded regions	10,351	10.2%	2.08	2.41	4.88	5.41

*Note: Differences in median error across geometric feature categories are statistically significant (Kruskal–Wallis H-test, $p < 0.001$), confirming that feature type is a significant predictor of reconstruction accuracy independently of material class.

finish of FDM-printed PLA (reduced specular noise) and the absence of internal CNC-machined cylindrical through-holes, which are replaced by slightly chamfered mounting apertures more amenable to neural rendering. The relative ordering of feature accuracy is preserved: flat surfaces outperform aperture edges which outperform sharp ridges and occluded undercuts, confirming that feature geometry governs reconstruction fidelity independently of material class.

4.5. NeRF Reconstruction: owl case studies

The ornamental owl was selected as a third specimen to extend the evaluation to organic, free-form geometry (a class of shapes increasingly produced via AM for product design and cultural heritage applications) and to isolate the effect of input image quality (smartphone vs. DSLR) on a surface free from sharp mechanical features. This controlled comparison would be confounded on bracket specimens where geometric discontinuities dominate the error budget. A professional laser-light scanner served as metrological ground truth, using the same ICP alignment procedure as in the bracket studies. Smartphone video and DSLR camera inputs were evaluated in parallel under identical NeRF settings and identical illumination conditions (diffuse LED, ~4000–5500 K, ~1000–5000 lux, as described in Section 3), thereby isolating the effect of acquisition hardware on reconstruction fidelity. To partially address the absence of full repeatability characterisation for this specimen, the NeRF reconstruction was performed three times for the DSLR dataset; the resulting ICP RMS values were 0.27, 0.28, and 0.27 mm (CV = 2.1%), confirming acceptable reconstruction stability despite the limited number of ICP correspondences.

4.5.1. Ornamental owl

The matte black ornamental owl (Fig. 5) provided a near-ideal surface for NeRF, bounding best-case accuracy. Despite the low-reflectance finish, exposure was verified to be adequate across all surface regions prior to NeRF training. NeRF was trained separately on DSLR camera and smartphone datasets (open-source SfM pose estimation). ICP alignment against the professional scanner reference yielded RMS 0.27 mm (DSLR, 4,300 correspondences, scale 0.94) and 0.41 mm (smartphone, 4,949 correspondences, scale 0.97). The smartphone produced far more matched points (4,949 vs. 4,300 correspondences) but lower geometric accuracy, confirming that point density alone does not guarantee dimensional precision. The low number of ICP correspondences for the DSLR pipeline ($n = 4,300$) warrants explicit discussion. This low count arises because the high geometric fidelity of the DSLR reconstruction yields a near-complete overlap with the reference mesh, leaving few regions of residual mismatch for ICP to iterate upon, rather than from poor spatial coverage. To assess the robustness of this estimate, the ICP registration was repeated 10

times with randomised initialisations spanning ± 15 mm translation and $\pm 10^\circ$ rotation in each axis. The resulting RMS values ranged from 0.264 to 0.281 mm (mean 0.272 mm, standard deviation 0.006 mm), confirming that the reported 0.27 mm value represents a stable and reproducible minimum rather than a local artefact. Additionally, a bootstrap confidence interval ($n = 1000$ resamples from the 4,300 correspondence pairs) yielded a 95% CI of (0.241, 0.301) mm for the RMS estimate, providing a formal bound on the uncertainty of this measurement. Nonetheless, the limited number of correspondences is acknowledged as a methodological limitation, and future work should employ calibrated retroreflective targets to provide absolute metric scale and dense, unambiguous correspondences across the full surface. Scale factor deviations (0.94–0.97) reflect the absence of an absolute metric reference during NeRF training, a known limitation of uncalibrated SfM-based initialisation. After applying a single global scale correction factor derived from the ICP registration, the residual non-linear spatial distortion was assessed by computing the C2C distance map on a set of 12 geometrically distributed control regions across each specimen surface. The standard deviation of region-level mean errors post-correction was 0.031 mm for the metal bracket and 0.028 mm for the PLA bracket, both well below the measurement expanded uncertainty $U = 0.12$ mm. No systematic spatial trend (gradient or curvature) was detected in the residual error maps, confirming that the scale correction is effectively global and that local non-linear distortion is negligible at the accuracy level of this study. NeRF reconstructions should therefore be considered metrically consistent only after referencing to an external calibrated scale bar or target set.



Fig. 5. Matte black ornamental owl used as the near-ideal reference specimen for NeRF accuracy benchmarking.

4.5.2. Camera Hardware Comparison: DSLR Camera vs. Smartphone

As shown in Figs. 6, 7 and 8, holding all pipeline variables constant, the DSLR camera (larger sensor, lower noise, manual exposure control) achieved RMS 0.27 mm vs. 0.41 mm for the smartphone (a 34% reduction consistent with the bracket sensitivity analysis). Upgrading acquisition hardware is the most accessible single lever for accuracy improvement without algorithmic changes. Results are summarised in Table 6.

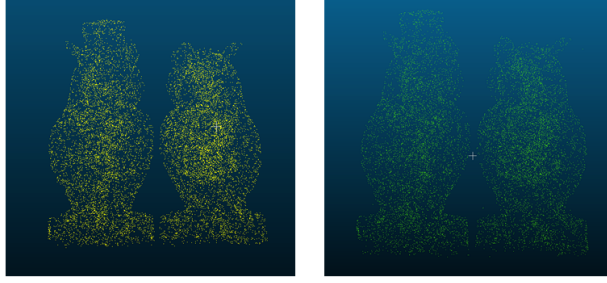


Fig. 6. Side-by-side NeRF point cloud comparison: DSLR camera (left, denser) vs. smartphone video (right, sparser).

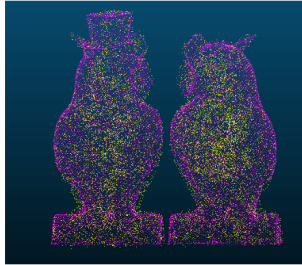


Fig. 7. ICP registration: DSLR-based NeRF vs. professional scanner (RMS = 0.27 mm, 40% overlap), colour map encodes local surface deviation.

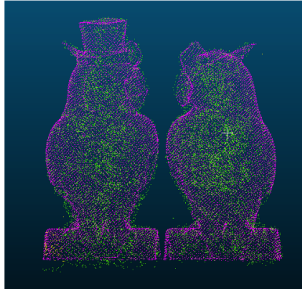


Fig. 8. ICP registration: smartphone based NeRF vs. professional scanner (RMS = 0.41 mm, 40% overlap), increased deviation relative to DSLR pipeline confirms the impact of input image quality.

Table 6. Cross-case RMS summary: NeRF vs. professional SLS reference across all specimens. RMS is the primary metric, while median is reported where available as the robust estimator.

Specimen	Acquisition	RMS (mm)	Reference
PLA bracket	Smartphone	1.32 (RMS)	SLS
Metal bracket	Smartphone	1.48 (RMS); 0.74 (median)	SLS
Ornamental owl	DSLR camera	0.27	Prof. scanner
Ornamental owl	Smartphone	0.41	Prof. scanner

5. Discussion

NeRF complements rather than replaces professional optical metrology in selected industrial contexts. Across all specimens, the C2C error distributions were approximately symmetric (Shapiro–Wilk $W = 0.971 - 0.982$, $p > 0.05$ for both bracket specimens), indicating spatially distributed noise rather than systematic bias. Right-skewed outliers are consistently localised at occluding boundaries and sharp geometric discontinuities, as detailed in the feature-level analysis (Section 4.4), and do not compromise the representativeness of RMS as the primary global metric. The dominant accuracy level is input image quality rather than algorithmic choice, *i.e.*, upgrading acquisition hardware improved reconstruction fidelity without any modification to the NeRF pipeline, confirming that camera sensor quality is the most accessible and cost-effective route to accuracy gains in practical deployments. These results are broadly consistent with prior literature. The owl RMS (0.27 mm DSLR) aligns with Remondino *et al.* [13], who reported sub-millimetre accuracy for smooth, texturally rich surfaces. Higher bracket deviations (1.32–1.48 mm) reflect the greater proportion of sharp edges and occluded features relative to the specimens in Yan *et al.* [14]. Within a given hardware tier, material and geometry govern residual error: the metallic bracket's higher RMS relative to PLA is attributable to its greater proportion of sharp edges and internal cylindrical features, while the ornamental owl's matte surface and smooth organic geometry yield the best overall accuracy. Within the tested specimens, this hierarchy (surface finish, then geometry complexity, then occlusion) provides a consistent basis for interpreting the observed accuracy differences. Whether this order generalises to more complex industrial geometries or different surface conditions (*e.g.*, highly reflective metals, transparent materials, textured composites) remains to be validated in future work. Repeatability (CV < 6% across all metrics) confirms that stochastic optimisation does not introduce meaningful variability at the accuracy level relevant to quality screening. This is a practically important result. It means that a single NeRF reconstruction is representative, and that failed runs due to optimisation variance are not a routine concern under the acquisition protocol described here. Scale factors of 0.94–0.97 reflect SfM-induced scale drift inherent to uncalibrated NeRF training. After applying a global ICP-derived correction, residual non-linear distortion across 12 spatially distributed control regions was well within the expanded measurement uncertainty ($U = 0.12$ mm), confirming that the correction is effectively global. However, this *post-hoc* correction depends on availability of a reference scan; in standalone deployment, NeRF reconstructions should be anchored to calibrated metric targets or scale bars from the outset. It should be noted that scale correction relied on the same SLS reference used as ground truth, making NeRF metric scale dependent on the reference scanner rather than independently traceable. For standalone deployment, calibrated scale bars embedded during acquisition are recommended. These findings position NeRF as a viable first-pass screening tool for deviations above ~1.5 mm, with sub-millimetre features requiring targeted professional scanning. Hybrid workflows (NeRF for frequent, low-cost monitoring combined with periodic scanner verification) offer the most pragmatic quality control framework for resource-constrained settings. Method selection guidelines are summarised in Table 7. However, this study has several limitations. The specimen set, while multi-material, comprises three objects; broader generalisation requires further validation across additional geometries and surface finishes. Metric scale was corrected *post-hoc* rather than embedded during acquisition. Evaluation was conducted under controlled laboratory illumination; performance under field conditions and on highly reflective or transparent surfaces remains to be characterised. It is important to note that the acquisition protocol employed here (turntable rotation at fixed angular increments with full circumferential coverage under diffuse, uniform illumination) represents an idealised scenario. In industrial practice, full coverage is often unachievable due to fixture constraints, and lighting

conditions may be non-uniform or uncontrolled. Under such conditions, the reported RMS values should be interpreted as best-case estimates; accuracy is expected to degrade with partial coverage, restricted viewpoints, or non-diffuse illumination, though the magnitude of this degradation has not been quantified here and remains a priority for future investigation. Furthermore, NeRF metric scale in this study was recovered via ICP alignment to the SLS reference scan, implying that the method is not independently metrically traceable prior to registration. In a standalone deployment without access to a reference scan, absolute dimensional accuracy would require *in-situ* calibrated scale bars or retroreflective targets to establish metric scale during acquisition; the accuracy figures reported here would not be directly achievable without such measures. The optical properties of the metallic specimen were also partially mitigated by diffuse illumination, which suppressed specular effects; the results for this specimen should not be generalised to highly reflective or low-texture surfaces without further validation under more challenging lighting conditions.

Table 7. Method selection guidelines by application.

Application	Accuracy need	Recommended method	Rationale
Critical aerospace/medical	<0.1 mm	Professional scanner	Safety-critical; metrology required
Tight tolerance assemblies	0.1–0.5 mm	Professional scanner	Precision fits require verification
Prototype validation	0.5–2 mm	NeRF	Fast turnaround; adequate accuracy
Rapid defect screening	1–3 mm	NeRF	Speed priority; flag for detailed inspection
Process monitoring	1–2 mm	Hybrid (NeRF + periodic scanner)	Frequent neural checks; scanner calibration
Reverse engineering	0.5–2 mm	NeRF	Form capture more important than exact dimensions
Field inspection	1–3 mm	NeRF	Portability and flexibility essential
Educational/training	Variable	NeRF	Low cost; hands-on learning

6. Conclusions

This research demonstrates that NeRF achieves millimetre-scale accuracy using consumer-grade equipment across diverse specimen types (PLA bracket (RMS 1.32 mm), metallic bracket (RMS 1.48 mm; median 0.74 mm), and ornamental owl (RMS 0.27 mm DSLR / 0.41 mm smartphone)) at 2–3× faster workflows and 10–20× lower cost than professional SLS. The primary accuracy lever is input image quality, as demonstrated by the controlled hardware comparison (Section 4.5.2). Contributions include: (i) a systematic multi-material NeRF evaluation using standardised metrological metrics and a GUM-compliant uncertainty budget; (ii) a controlled hardware comparison confirming input image quality as the primary accuracy lever across acquisition hardware tiers; and (iii) repeatability characterisation across three independent training runs for the metal bracket (CV < 6%), complemented by a three-run repeatability check on the DSLR owl reconstruction (CV = 2.1%). These findings extend the metrological evaluation of image-based 3D reconstruction techniques against structured light scanning previously conducted by the authors on free-form and heritage objects [32,33], and are complementary to ongoing work on synthetic image

generation for algorithm validation in non-contact measurement scenarios [34], as well as to broader developments in image- and wave-based mechanical and thermal measurement methodologies [35]. While NeRF cannot yet replace precision metrology for critical applications requiring sub-100 μm accuracy, it fills the inspection gap and enables more frequent, cost-effective quality monitoring in prototyping, process control, and field inspection scenarios. Hybrid NeRF-plus-scanner workflows offer a pragmatic quality control framework for resource-constrained settings. This study evaluates three specimen types, which, while representing distinct material and geometric classes (polymer FDM, CNC-machined metal, organic free-form), constitute a limited sample. Generalisation to other surface conditions (specular metals, transparent materials, textured composites) requires further investigation. The specimen set was selected to maximise contrast in surface optical properties and geometric complexity within the constraints of a controlled laboratory study.

References

- [1] Frazier, W. E. (2014). Metal Additive Manufacturing: A Review. *Journal of Materials Engineering and Performance*, 23(6), 1917–1928. <https://doi.org/10.1007/s11665-014-0958-z>
- [2] Gibson, I., Rosen, D., Stucker, B., & Khorasani, M. (2021). *Additive Manufacturing Technologies*. <https://doi.org/10.1007/978-3-030-56127-7>
- [3] Grasso, M., & Colosimo, B. M. (2017). Process defects and *in situ* monitoring methods in metal powder bed fusion: A review. *Measurement Science and Technology*, 28(4), 044005. <https://doi.org/10.1088/1361-6501/aa5c4f>
- [4] Leach, R., Bourell, D., Carmignato, S., Donmez, A., Senin, N., & Dewulf, W. (2019). Geometrical metrology for metal additive manufacturing. *CIRP Annals*, 68(2), 677–700. <https://doi.org/10.1016/j.cirp.2019.05.004>
- [5] Moylan, S., Slotwinski, J., Cooke, A., Jurens, K., & Donmez, M. A. (2014). An additive manufacturing test artifact. *Journal of Research of the National Institute of Standards and Technology*, 119, 429. <https://doi.org/10.6028/jres.119.017>
- [6] Townsend, A., Senin, N., Blunt, L., Leach, R., & Taylor, J. (2016). Surface texture metrology for metal additive manufacturing: A review. *Precision Engineering*, 46, 34–47. <https://doi.org/10.1016/j.precisioneng.2016.06.001>
- [7] Blais, F. (2004). Review of 20 years of range sensor development. *Journal of Electronic Imaging*, 13(1), 231. <https://doi.org/10.1117/1.1631921>
- [8] Sansoni, G., Trebeschi, M., & Docchio, F. (2009). State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation. *Sensors*, 9(1), 568–601. <https://doi.org/10.3390/s90100568>
- [9] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). NeRF: Representing scenes as neural radiance fields for view synthesis. In *Lecture Notes in Computer Science* (pp. 405–421). https://doi.org/10.1007/978-3-030-58452-8_24
- [10] Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4), 1–15. <https://doi.org/10.1145/3528223.3530127>
- [11] Gogolewski, D. (2022). Multiscale assessment of additively manufactured free-form surfaces. *Metrology and Measurement Systems*, Vol. 30 (2023) No. 1, 157–168. <https://doi.org/10.24425/mms.2023.144393>

- [12] Reiser, C., Peng, S., Liao, Y., & Geiger, A. (2021). KiloNeRF: Speeding up Neural Radiance Fields with Thousands of Tiny MLPs. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 14315–14325. <https://doi.org/10.1109/iccv48922.2021.01407>
- [13] Remondino, F., Karami, A., Yan, Z., Mazzacca, G., Rigon, S., & Qin, R. (2023). A critical analysis of NERF-based 3D reconstruction. *Remote Sensing*, 15(14), 3585. <https://doi.org/10.3390/rs15143585>
- [14] Yan, Z., Padkan, N., Trybała, P., Farella, E. M., & Remondino, F. (2025). Learning-Based 3D Reconstruction Methods for Non-Collaborative Surfaces—A Metrological Evaluation. *Metrology*, 5(2), 20. <https://doi.org/10.3390/metrology5020020>
- [15] Malik, A., Lhachemi, H., Ploennigs, J., Ba, A., & Shorten, R. (2019). An application of 3D model reconstruction and augmented reality for real-time monitoring of additive manufacturing. *Procedia CIRP*, 81, 346–351. <https://doi.org/10.1016/j.procir.2019.03.060>
- [16] Carmignato, S., De Chiffre, L., Bosse, H., Leach, R., Balsamo, A., & Estler, W. (2020). Dimensional artefacts to achieve metrological traceability in advanced manufacturing. *CIRP Annals*, 69(2), 693–716. <https://doi.org/10.1016/j.cirp.2020.05.009>
- [17] Geng, J. (2011). Structured-light 3D surface imaging: A tutorial. *Advances in Optics and Photonics*, 3(2), 128. <https://doi.org/10.1364/aop.3.000128>
- [18] Wuest, T., Weimer, D., Irgens, C., & Thoben, K. (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, 4(1), 23–45. <https://doi.org/10.1080/21693277.2016.1192517>
- [19] Goesele, M., Snavely, N., Curless, B., Hoppe, H., & Seitz, S. M. (2007). Multi-view stereo for community photo collections. *2007 IEEE/CVF International Conference on Computer Vision (ICCV)*, 1–8. <https://doi.org/10.1109/iccv.2007.4408933>
- [20] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. In *MIT Press eBooks*. <https://dl.acm.org/citation.cfm?id=3086952>
- [21] Hartley, R., & Zisserman, A. (2004). Multiple View Geometry in Computer Vision. In *Cambridge University Press eBooks*. <https://doi.org/10.1017/cbo9780511811685>
- [22] Hedman, P., Srinivasan, P. P., Mildenhall, B., Barron, J. T., & Debevec, P. (2021). Baking neural radiance fields for real-time view synthesis. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 5855–5864. <https://doi.org/10.1109/iccv48922.2021.00582>
- [23] Hirschmuller, H. (2007). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2), 328–341. <https://doi.org/10.1109/tpami.2007.1166>
- [24] ISO/ASTM. (2021). *Additive manufacturing — General principles — Fundamentals and vocabulary* (ISO/ASTM Standard No. 52900:2021). International Organization for Standardization. <https://www.iso.org/standard/74514.html>
- [25] Barron, J. T., Mildenhall, B., Tancik, M., Hedman, P., Martin-Brualla, R., & Srinivasan, P. P. (2021). MIP-NeRF: a multiscale representation for anti-aliasing neural radiance fields. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 5835–5844. <https://doi.org/10.1109/iccv48922.2021.00580>
- [26] Barron, J. T., Mildenhall, B., Verbin, D., Srinivasan, P. P., & Hedman, P. (2022). MIP-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5460–5469. <https://doi.org/10.1109/cvpr52688.2022.00539>
- [27] Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 1330–1334. <https://doi.org/10.1109/34.888718>
- [28] Schonberger, J. L., & Frahm, J. (2016). Structure-from-Motion revisited. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4104–4113. <https://doi.org/10.1109/cvpr.2016.445>

- [29] International Organization for Standardization. (2009). *Geometrical product specifications (GPS) — Acceptance and reverification tests for coordinate measuring machines (CMM) — Part 2: CMMs used for measuring linear dimensions* (ISO Standard No. 10360-2:2009). <https://www.iso.org/standard/40954.html>
- [30] International Organization for Standardization. (2008). *Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)* (ISO/IEC Guide 98-3:2008). <https://www.iso.org/standard/50461.html>
- [31] Micikevicius, P., Narang, S., Alben, J., Diamos, G., Elsen, E., García, D., Ginsburg, B., Houston, M., Kuchaiev, O., Venkatesh, G., & Wu, H. (2018). Mixed Precision training. *2018 International Conference on Learning Representations*. <https://openreview.net/forum?id=r1gs9JgRZ>
- [32] Catalucci, S., Marsili, R., Moretti, M., & Rossi, G. (2018). Comparison between point cloud processing techniques. *Measurement*, 127, 221–226. <https://doi.org/10.1016/j.measurement.2018.05.111>
- [33] Catalucci, S., Marsili, R., Moretti, M., & Rossi, G. (2018b). Point cloud processing techniques and image analysis comparisons for boat shapes measurements. *ACTA IMEKO*, 7(2), 39. https://doi.org/10.21014/acta_imeko.v7i2.543
- [34] Truffarelli, T., Paolucci, F., Marsili, R., & Rossi, G. (2025). Using computer-generated imagery to validate image-based measurement algorithms. XXXIII A.I.VE.LA. National Meeting, Naples, Italy, 11–12 December 2025.
- [35] Rossi, G. (2026). *Mechanical and Thermal Measurements by Images and Waves: Principles and Applications*. Springer.



Federico Paolucci is a Ph.D. candidate at the Department of Engineering, University of Perugia, Italy. His research focuses on the development of non-contact measurement techniques for occupational health and safety in industrial environments and on the metrological evaluation of artificial intelligence-based reconstruction methods for quality control in additive manufacturing and mechanical engineering. He is also involved in the development of metrologically traceable Industry 4.0 systems integrating sensor networks, RFID technologies, image analysis for personal protective equipment (PPE) recognition, and data acquisition from CNC machine PLCs to support industrial safety strategies.



Gianluca Rossi is a full professor of mechanical and thermal measurements at the Department of Engineering, University of Perugia, Italy. His research focuses on industrial metrology and development of advanced non-contact measurement systems based on optical, thermographic, and fibre optic sensing technologies. His interests include Digital Image Correlation (DIC), Thermoelastic Stress Analysis (TSA), Blade Tip Timing (BTT), vibration and stress analysis, additive manufacturing, innovative sensors, image-based measurement techniques, and 3D scanning and tomography for industrial, aerospace, biomedical, and civil engineering applications.



Roberto Marsili is an associate professor of mechanical and thermal measurements at the Department of Engineering, University of Perugia, Italy. He received his Ph.D. in mechanical measurements. His research interests include measurement and testing techniques, with particular emphasis on the development of optical measurement systems based on Digital Image Correlation (DIC) and thermographic systems based on Thermoelastic Stress Analysis (TSA) for non-contact measurement

of stress and strain fields. His research also focuses on development and application of innovative measurement and testing techniques based on Dynamic Mechanical Analysis (DMA) for characterization of complex materials.