Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

# Computer Vision and Sensor Measurements for Crack Monitoring of Reinforced Concrete Beams: A Critical Review

Parita Patel
M.Tech Scholar (Structural Engineering)
Department of Civil Engineering
Chhotubhai Gopalbhai Patel Institute of Technology
Uka Tarsadia University, Bardoli, Gujarat, India

Maulik Kakadiya
Assistant Professor
Department of Civil Engineering
Chhotubhai Gopalbhai Patel Institute of Technology
Uka Tarsadia University, Bardoli, Gujarat, India

Palak Trivedi
Assistant Professor
Department of Civil Engineering
Chhotubhai Gopalbhai Patel Institute of Technology
Uka Tarsadia University, Bardoli, Gujarat, India

***Abstract*—Visible cracking is the earliest and most accessible evidence of distress in reinforced concrete beams, yet in most laboratories and inspections it is still recorded by hand-marking cracks and reading their widths with a comparator, a process that is slow, depends on the inspector and is hard to relate to the load carried by the member. Computer vision and deep learning can now locate and outline cracks in ordinary photographs, while inexpensive sensors record load, deflection and strain during a test. This paper presents a narrative critical review of thirty sources on vision-based and sensor-based crack monitoring, with emphasis on reinforced concrete beams. The literature is grouped into five strands: monitoring principles, classical image processing, deep-learning classification, pixel-level segmentation, and crack-width measurement on loaded members. The comparison shows that learning-based segmentation consistently suppresses the background noise that limits edge detectors and thresholding, and that encoder-decoder networks of the U-Net family give the boundary detail needed for width estimation. Reported scores, however, are seldom comparable, because studies differ in task definition, image resolution, reference measurement and data splitting. Few studies place image-derived crack metrics beside measured deflection or strain at matching load stages, and pixel size is rarely treated as a design parameter of the test. Six research gaps are identified, together with the elements of a reproducible laboratory protocol that could address them. The review is intended to help structural engineers plan vision-assisted beam tests whose outputs can be read alongside conventional instrumentation.**

***Keywords*—Reinforced concrete beam, crack detection, crack width, computer vision, deep learning, U-Net, semantic segmentation, structural health monitoring***

## I. INTRODUCTION

Reinforced concrete (RC) beams are designed on the assumption that concrete in the tension zone will crack. Because the tensile strength of concrete is low, flexural cracks open in service at moment levels well below the design capacity, and their number and width grow as the reinforcement takes up more stress. Indian practice limits the surface width of such cracks to 0.3 mm in general and to 0.2 mm where cracking may harm durability, and IS 456 gives an expression for estimating the design surface crack width [1]. Crack width is therefore both a serviceability criterion and an index of how far a member has been strained.

In laboratories and in field inspection alike, cracks are usually traced by eye, marked on the surface and measured with a crack comparator or a hand microscope. The method is simple but has well-known weaknesses: it is slow, its results depend on the observer, repeated surveys are hard to compare, and the readings are not tied to the load or deformation of the member at the moment they were taken [2], [3]. Several reviews have traced the effort to automate the task with image processing [4], [5].

Structural health monitoring (SHM) offers a broader framework in which measured response is converted into damage-sensitive features and then into a judgement about condition [6]. Much SHM research has relied on global vibration measurements [7], which are informative for whole structures but say little about where a beam has cracked. Cameras supply the missing local, spatially dense information, and vision-based SHM is now a recognised field with its own review literature [3], [8], [9]. Deep learning has accelerated both strands [10].

Existing reviews mostly survey crack-detection algorithms across pavements, bridge decks and building surfaces [2], [4], [5]. They seldom ask a narrower question that matters to structural engineers: what is required when an image is used as a measuring instrument in a beam test, so that crack metrics can be set beside the readings of load cells, displacement transducers and strain gauges? This review takes that perspective. Its contributions are as follows.

- It separates the image-analysis tasks of classification, segmentation and measurement and shows why their reported scores cannot be compared directly.
- It compares studies that measure crack width from images or relate crack images to the structural response of loaded members, with attention to resolution, reference measurement and data splitting.
- It distinguishes conclusions that are well supported by several independent studies from those that rest on a single dataset or test.



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

- It derives six research gaps and relates them to the design of a reproducible beam-testing protocol being developed in the first author's ongoing M.Tech dissertation.

Section II describes how the literature was assembled. Section III sets out the structural and measurement background. Section IV reviews the literature by theme, Section V compares studies critically and Section VI synthesises the evidence. Research gaps, engineering implications and conclusions follow in Sections VII to IX.

## II. REVIEW SCOPE AND METHOD

This is a narrative critical review. No exhaustive database search or PRISMA-type screening was carried out, and the paper does not claim to cover every published method. Four research questions guided the selection and comparison of sources:

- RQ1: Which image-based methods have been used to detect and delineate cracks on concrete and masonry surfaces, and how do they compare?
- RQ2: How accurately has crack width been obtained from images, and against what reference measurement?
- RQ3: To what extent have crack images been combined with the measured structural response of loaded members?
- RQ4: Which methodological factors limit the comparability and transferability of reported results?

The starting point was the reference list compiled for the first author's dissertation, each entry of which was checked against the publisher or repository record. It was supplemented in October 2026 by targeted searches of publisher sites and open repositories using combinations of the terms concrete crack, segmentation, U-Net, crack width, digital image correlation, reinforced concrete beam and structural health monitoring. Sources were included if they were peer-reviewed journal or conference publications (or established standards and software used in the field), dealt with cracks on concrete or masonry surfaces, described both their method and its evaluation, and had bibliographic details that could be verified. Pavement-only studies, unreviewed preprints and studies of subsurface or steel fatigue cracking were excluded.

Each study was coded by task, material and specimen, data source, evaluation metric, reference used for crack width, and whether structural response was measured. Numerical findings are quoted as the original authors reported them; they have not been re-analysed, and figures obtained on different datasets are not treated as directly comparable.

## III. CRACKING IN RC BEAMS AND MONITORING REQUIREMENTS

### *A. Flexural Cracking and Its Significance*

In a simply supported beam under two-point loading, the first flexural cracks form in the constant-moment region once the stress in the extreme tension fibre exceeds the flexural tensile strength of concrete. Further loading adds cracks at roughly regular spacing and widens existing ones in proportion to the strain in the tension steel, modified by the cover and bar spacing; the crack-width expression in IS 456 reflects these variables [1]. A crack pattern therefore records where the tensile capacity of concrete has been exceeded, a crack offers a path for moisture and chlorides to the reinforcement, and wide cracks impair appearance and function. A surface crack width is not, however, a direct measure of safety: the same width can arise at different steel stresses in members with different cover. It is useful to separate four tasks: detecting cracks in an image, measuring their geometry, monitoring the structural response and judging structural safety. Image analysis can contribute to the first two, sensors to the third, and only engineering assessment to the fourth.

### *B. Sensor-Based Monitoring*

Sensor-based SHM treats damage identification as a statistical pattern-recognition problem: measurements are acquired, features sensitive to damage are extracted, and a model classifies the condition of the structure [6]. Over the last decade most of this work has used vibration response [7]. In a laboratory beam test the relevant measurements are simpler: applied load from a load cell or calibrated jack, deflection from dial gauges or linear variable differential transformers (LVDTs), and strain from electrical resistance gauges. These readings are precise and physically calibrated, but they are sparse; they reveal that stiffness has fallen without showing where or how the beam has cracked.

Vision fills that gap. Spencer et al. distinguish inspection applications, which identify and characterise visible damage, from monitoring applications, which measure displacement and strain [3], and Dong and Catbas review vision-based SHM at both local and global levels [8]. Ye et al. draw attention to systematic error sources in machine-vision measurement and to the value of integrating cameras with other sensors [9]. Fig. 1 shows how the two streams, together with a manual reference, can be joined in a stage-matched record.

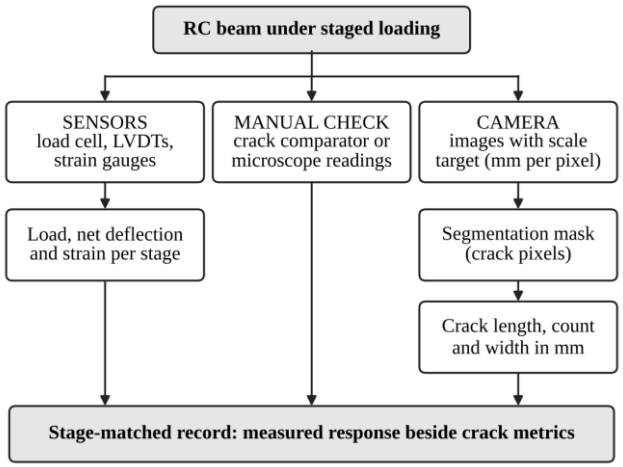



Fig. 1. Integration of sensor readings, manual crack-width checks and image-derived crack metrics in a stage-matched record for a laboratory RC beam test (schematic prepared for this review).

### *C. From Image to Measurement*

The term crack detection hides several distinct tasks, summarised in Table I. Only semantic segmentation, which labels each pixel as crack or background, yields the crack length and width needed for engineering interpretation; a classifier with high patch accuracy may still be unable to measure anything. Converting a mask to physical units requires the



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

image scale $s$, obtained from a target of known size placed on the specimen surface:

$$s = L_{\text{ref}} / n_{\text{px}} \quad (1)$$

where $L_{\text{ref}}$ is the known length on the target (mm) and $n_{\text{px}}$ the number of pixels spanning it. A common way of extracting width is to thin the mask to a one-pixel skeleton and evaluate a distance transform; at skeleton pixel $i$ the local width is

$$w_i = 2\, s\, D(p_i) \quad (2)$$

where $D(p_i)$ is the distance in pixels from the skeleton pixel to the nearest background pixel. Skeleton-based measurement of this kind underlies the width estimates of Yang et al. [11]. Because a mask boundary cannot be located more precisely than about one pixel, the relative uncertainty of a width estimate is of the order of $s/w$. Fig. 2 plots this simple geometric ratio. At 0.075 mm per pixel, a typical value for a 12-megapixel image covering 300 mm of beam, a 0.15 mm crack spans only two pixels and a one-pixel error changes its width by half; near the 0.2 mm limit of IS 456 the uncertainty is still about 38 per cent. Close-up images with pixel sizes of 0.03 mm or finer reduce it to roughly 15 per cent.

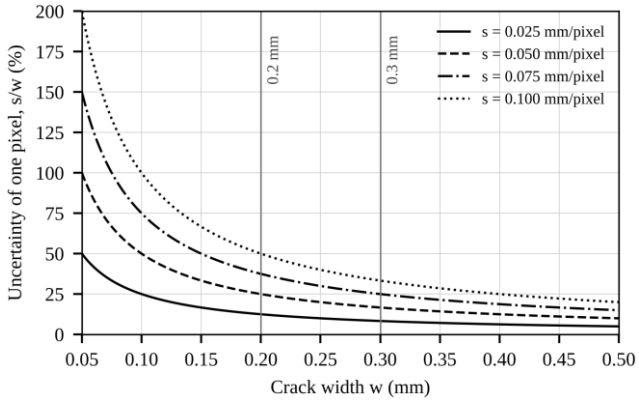



Fig. 2. Relative uncertainty of a width estimate caused by a one-pixel boundary error, s/w, for four pixel sizes s; vertical lines mark the 0.2 mm and 0.3 mm limits of IS 456. Geometric illustration, not experimental data.

TABLE I.
IMAGE-ANALYSIS TASKS RELEVANT TO CRACK MONITORING

| Task | Output | Typical metric | Gives width? |
|---|---|---|---|
| Image or patch classification | Crack / no crack per patch | Accuracy, precision, recall, F1 (per patch) | No |
| Object detection | Bounding boxes | Box overlap, average precision | No |
| Semantic segmentation | Crack mask per pixel | Pixel precision, recall, F1 (Dice), IoU | Yes, after scaling |
| Measurement | Length, count, width in mm | Error against a physical reference | Yes |

## IV. THEMATIC REVIEW OF EXISTING RESEARCH

### *A. Classical Image Processing*

The earliest automated methods used global or adaptive thresholding, of which Otsu's histogram-based method [12] remains a common choice, together with edge detectors such as Sobel and Canny, morphological filtering, and ridge or vesselness filters such as the multiscale Hessian filter of Frangi et al. [13], which responds to thin elongated structures. Koch et al. synthesised the achievements and limitations of these approaches for concrete and asphalt infrastructure [2]. Mohan and Poobal compared some fifty image-processing studies by technique, objective, accuracy and dataset [4], and Munawar et al. compared thirty studies from the preceding decade [5].

The clearest quantitative evidence comes from Dorafshan et al., who applied six edge detectors and a deep convolutional neural network (DCNN) to the same set of 19 high-resolution concrete images divided into 3,420 sub-images [14]. The edge detectors found roughly 53 to 79 per cent of cracked pixels but left considerable residual noise, while the DCNN detected cracks down to about 0.08 mm when fully trained and about 0.04 mm with transfer learning; a hybrid of the two reduced noise substantially. The general picture is consistent across the reviews: classical methods need no training data and are transparent and cheap, but their parameters must be tuned to each surface, and they respond to formwork marks, stains, shadows and texture as readily as to cracks. Today they serve best as baselines and pre-processing steps.

### *B. Deep Learning for Crack Classification*

Cha et al. showed that a convolutional neural network (CNN) could classify concrete image patches as cracked or intact without hand-crafted features, and that it outperformed Canny and Sobel detectors on their images [15]. Public datasets followed. SDNET2018 contains more than 56,000 images of cracked and uncracked bridge decks, walls and pavements, with crack widths from about 0.06 mm to 25 mm and realistic obstructions such as shadows and surface roughness [16]; its labels, however, are assigned per image rather than per pixel. Dung and Anh reported classification accuracy of at least 97.8 per cent on a public set of 40,000 patches using networks pretrained on general images [17].

Such accuracies should be read with care. Patch-level scores depend on the balance between cracked and intact patches and on how difficult the negative examples are, and a correct label says nothing about where the crack lies within the patch or how wide it is. For monitoring a beam, classification is useful mainly for screening large image sets and for pretraining feature extractors.

### *C. Pixel-Level Segmentation*

Fully convolutional networks (FCNs) removed the need for sliding windows by predicting a label for every pixel. Yang et al. trained an FCN to segment several crack types and reduced the predictions to skeletons from which crack topology, length and width were computed; they reported precision of 81.73 per cent, recall of 78.97 per cent and an F1 score of 79.95 per cent [11]. Dung and Anh trained an encoder–decoder FCN with a VGG16 encoder on 500 annotated images, reporting average precision of about 90 per cent and reasonable detection on video frames from a cyclic loading test [17]. Liu et al. introduced DeepCrack, a hierarchical network trained with deep supervision, and published a benchmark of 537 annotated images on which they obtained a mean intersection over union (IoU) of 85.9 and a best F-score of 86.5 at about 0.1 s per image [18].

The U-Net of Ronneberger et al. combines a contracting encoder with an expanding decoder and transfers high-



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

resolution features through skip connections [19], which suits thin structures such as cracks. When annotated data are scarce, the encoder is usually initialised with weights from a network such as ResNet [20] trained on general images. Dais et al. used this transfer-learning approach on masonry, reaching 95.3 per cent accuracy at patch level and an F1 score of 79.6 per cent at pixel level [21]. Ariyaratne et al. compared U-Net, Attention U-Net, DeepLabV3+ and YOLOv8-seg on 2,000 annotated images under identical training conditions [22]. The two U-Net variants gave the best IoU (0.741 and 0.740) and the most reliable widths, whereas the instance-segmentation detector was markedly less accurate. Open-source libraries now make such encoder–decoder models available with a few lines of code [23].

The studies agree that architectures preserving fine spatial detail delineate cracks better than detectors that return boxes or coarse masks. They disagree, or cannot be compared, on absolute performance: an F1 of about 0.8 in one study and a mean IoU of 85.9 in another reflect different images, annotation conventions and metric definitions rather than a ranking of methods.

### *D. Crack-Width Quantification*

Measuring width is harder than finding a crack. Jahanshahi and Masri proposed a photogrammetric, non-contact methodology for accurate crack-width measurement aimed at automated safety evaluation [24]. In the skeleton-based approach of Yang et al., relative width errors ranged from −13.27 to +24.01 per cent, while length errors were much larger, from −48.03 to +177.79 per cent [11]. On images from bending tests of RC beams, Qiao et al. used a U-Net improved with an inception module and atrous spatial pyramid pooling and reported an average relative width error of 13.2 per cent, lower than that of a commercial crack-measurement system [25]. Carrasco et al. combined anisotropic smoothing, segmentation and k-means-stabilised centre points; on fibre-reinforced earthen materials they detected cracks as narrow as 0.15 mm and found no statistically significant difference from 330 manual measurements, while working at least twenty times faster [26]. Ariyaratne et al. validated predicted widths against laboratory-tested RC beams [22].

These results are encouraging but heterogeneous. The reference against which width is judged varies from manual image measurement to commercial systems and physical readings; relative errors are inflated for narrow cracks, where a single pixel is a large fraction of the width; and image resolution is not always stated, so that the same relative error may correspond to very different absolute errors in millimetres.

### *E. Vision on Loaded Members and Link to Structural Response*

A smaller body of work uses crack images from members tested under load. Davoudi et al. developed data-driven, vision-based inspection of RC beams and slabs that estimates damage and load from crack features [27]. Ding et al. extracted crack distributions with a U-Net having an EfficientNet-B7 encoder, related them to strain fields from finite-element analysis, and estimated the applied load with a mean absolute percentage error of 10.98 per cent over 25 load conditions of seven beams; crack width was left for future work [28]. Both studies show that crack morphology carries information about load level, but both infer load through a model rather than reading it beside the crack metrics.

Digital image correlation (DIC) provides a different bridge between images and response. Słoński and Tekieli monitored precast post-tensioned crane beams in three-point bending with three synchronised cameras [29]. DIC revealed the first midspan crack at about 287 kN before it was visible to the eye, and DIC deflections agreed with LVDT readings to within 1 mm; a Faster R-CNN detector located the visible cracks, although no quantitative detection metrics were reported and the evaluation rested on one beam. Rezaie et al. compared crack segmentation by thresholding DIC principal-strain maps with a deep network on stone masonry walls and found the network about 47 per cent more precise and about 34 per cent better in Dice coefficient, at the cost of about 4 per cent lower sensitivity [30].

## V. CRITICAL COMPARATIVE ANALYSIS

Table II summarises the principal studies against the criteria used in this review. Entries are taken from the original publications; NR marks information not reported in the source consulted, and no values have been estimated to fill gaps.

TABLE II.
COMPARISON OF REVIEWED STUDIES ON IMAGE-BASED CRACK ANALYSIS

| Ref. | Material / specimen | Task | Method | Data | Reported finding | Response measured |
|---|---|---|---|---|---|---|
| [14] | Concrete surfaces | Classification, edge detection | Six edge detectors; AlexNet DCNN (three modes) | 19 images; 3,420 sub-images | Edge detectors 53–79% of cracked pixels; DCNN found finer cracks | No |
| [15] | Concrete surfaces | Patch classification | CNN vs Canny and Sobel | Concrete images | CNN outperformed edge detectors | No |
| [16] | Decks, walls, pavements | Dataset | Image-level labels | > 56,000 images | Crack widths 0.06–25 mm, with obstructions | No |
| [17] | Concrete | Classification; segmentation | Pretrained CNNs; VGG16 FCN | 40,000 patches; 500 masks | Accuracy ≥ 97.8%; AP about 90% | Video of cyclic test; no sensor link reported |
| [11] | Concrete | Segmentation, measurement | FCN with skeletonisation | NR | F1 79.95%; width error −13.27 to +24.01% | No |
| [18] | Mixed scenes | Segmentation | DeepCrack | 537 annotated images | Mean IoU 85.9; best F 86.5 | No |
| [21] | Masonry | Classification; segmentation | CNNs with transfer learning | Masonry photographs | Patch accuracy 95.3%; pixel F1 79.6% | No |



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

| Ref. | Material / specimen | Task | Method | Data | Reported finding | Response measured |
|---|---|---|---|---|---|---|
| [26] | Fibre-reinforced earthen material | Width measurement | Smoothing, segmentation, k-means | 330 manual readings | Cracks ≥ 0.15 mm; no significant difference from manual | No |
| [25] | RC beams in bending | Segmentation, width | Improved U-Net (inception, ASPP) | Beam test images | Mean relative width error 13.2% | NR |
| [22] | Concrete; lab RC beams | Segmentation, width | U-Net, Att. U-Net, DeepLabV3+, YOLOv8-seg | 2,000 annotated images | IoU 0.741 (U-Net); YOLOv8-seg less accurate | NR |
| [28] | RC beams | Segmentation for load estimation | U-Net (EfficientNet-B7) with FE strain fields | 7 beams; 25 load conditions | Load MAPE 10.98%; width not addressed | Load, via FE model |
| [27] | RC beams and slabs | Damage and load estimation | Data-driven crack features | NR | Load and damage estimated from crack patterns | Load |
| [29] | Precast PT crane beams | Detection; DIC | Faster R-CNN; 2D DIC | 1,058 training images; one beam | First crack about 287 kN by DIC; DIC within 1 mm of LVDT | Load, LVDT |
| [30] | Stone masonry walls | Segmentation | DIC strain threshold vs DCNN | Two test campaigns | DCNN precision +47%, Dice +34%, sensitivity −4% | DIC fields |

### *A. Task Definition and Metrics*

Scores in Table II belong to four different tasks. Patch accuracy above 95 per cent and pixel F1 near 80 per cent can describe equally competent models, because pixel metrics penalise every misplaced boundary pixel on a crack only a few pixels wide. Even within segmentation, some authors report IoU and others Dice or F1, at the level of individual images or of the whole set, with or without a tolerance band around the annotated crack. Without a common protocol, cross-study rankings are not meaningful, and a reader should compare methods only when they have been evaluated on the same images, as in [14], [22], [30].

### *B. Image Resolution and Measurable Width*

Detection and measurement have different resolution requirements. A network can often detect a crack narrower than one pixel because the crack darkens a line of pixels, as the 0.04 mm detection limit reported in [14] illustrates, but it cannot measure that crack's width. Fig. 2 shows that, at the pixel sizes common in overview photographs, widths near code limits carry uncertainties of tens of per cent. Few of the studies reviewed state pixel size alongside width error, which makes it impossible to judge whether a reported error reflects the algorithm or the camera.

### *C. Reference Measurements for Width*

Width errors are only as good as their reference. The reviewed studies used manual measurement [26], a commercial measurement system [25], laboratory beams [22] or DIC [29], [30]. A crack comparator gives a physical reading but has its own resolution and observer variability; DIC gives full-field displacement but depends on speckle quality and subset size. Studies rarely report the repeatability of the reference, so the scatter attributed to the algorithm may partly belong to the reference.

### *D. Data Splitting and Domain Shift*

Images taken from one specimen, one loading stage or one capture session are highly similar. If tiles from such images are split at random between training and test sets, nearly identical views appear in both and the test score overstates performance on a new member. The records reviewed seldom state whether training and test images came from different specimens. Domain shift is a related problem: models trained on public datasets of decks, walls and pavements [16], [18] meet different textures, lighting and crack widths on a whitewashed laboratory beam, which is why fine-tuning on specimen images followed by testing on a withheld specimen is preferable.

### *E. Computational Resources and Reproducibility*

Reported inference times, such as 0.1 s per image for DeepCrack [18], and the observation that end-to-end FCNs train much faster than earlier patch-based networks [11], suggest that segmentation is no longer computationally prohibitive. Pretrained encoders and open libraries [20], [23] allow a U-Net to be fine-tuned on modest hardware. Few papers, however, release their annotated beam images or code, which limits independent verification.

## VI. SYNTHESIS OF RESEARCH FINDINGS

Table III grades the main findings by the breadth of their support. The strongest conclusion is that learning-based methods suppress background noise better than classical image processing; it is supported on concrete, masonry and DIC data by studies that compared methods on identical images [14], [15], [30]. Moderately supported is the advantage of encoder–decoder architectures for thin cracks and width estimation, which follows from the design of U-Net [19] and from direct comparisons [22], but rests on few head-to-head studies. The view that crack morphology carries information about load is supported by two independent beam studies [27], [28], though both depend on a model linking cracks to load.

More tentative are the magnitude of width error achievable from images, reported as roughly 13 to 24 per cent in relative terms [11], [25] but with different references and unstated resolution, and the ability of DIC to reveal cracking before it becomes visible, observed in a single full-scale test [29]. These conclusions depend on particular specimens, cameras and reference methods and should not be generalised without replication.

TABLE III.
STRENGTH OF EVIDENCE FOR THE PRINCIPAL FINDINGS

| Finding | Support | Assessment |
|---|---|---|
| Learning-based segmentation suppresses background noise better than edge detection or thresholding | [14], [15], [30] | Strong; consistent across materials |



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

| Finding | Support | Assessment |
|---|---|---|
| Encoder–decoder networks suit thin cracks and width estimation | [11], [19], [22], [25] | Moderate; few direct comparisons |
| Crack morphology carries information on load level | [27], [28] | Moderate; model-dependent |
| Relative width error of about 13–24% is achievable | [11], [25] | Tentative; references and resolution differ |
| DIC can reveal cracking before it is visible | [29] | Tentative; single test |
| Patch accuracy does not indicate ability to measure | [15]–[17] | Strong; follows from task definition |

## VII. RESEARCH GAPS AND FUTURE DIRECTIONS

The comparison points to six gaps, listed in Table IV with the reason each matters and a possible way to close it. They are opportunities identified from the sources reviewed, not claims that no study has addressed them.

TABLE IV.
RESEARCH GAPS AND POSSIBLE DIRECTIONS

| Gap | Why it matters | Possible direction |
|---|---|---|
| G1. Few stage-matched records of crack metrics with load, deflection and strain | Crack growth cannot be interpreted without the response that caused it | Record sensor readings, images and manual widths at the same hold points |
| G2. Image resolution seldom treated as a design parameter | Width error is bounded by pixel size (Fig. 2) | Plan pixel size from expected widths; report mm per pixel |
| G3. Inconsistent or undocumented width references | Algorithm error and reference error are confounded | Use comparator or microscope readings at marked stations, with repeatability |
| G4. Random splits between training and test images | Leakage overstates performance on new members | Split by specimen; test once on a withheld beam |
| G5. Non-standard reporting of metrics | Results cannot be compared across studies | Report strict and tolerance-based pixel metrics and width error in mm |
| G6. Limited release of data and code | Results cannot be verified or reused | Use open libraries; publish masks, scale data and scripts |

The first author's ongoing M.Tech dissertation is being designed around gaps G1 to G4. The proposed programme comprises three nominally identical RC beams (150 mm × 200 mm × 1400 mm, M25 concrete and Fe500 reinforcement) tested in four-point bending in twelve load stages. At each stage, load and mid-span deflection (and strain where instruments are available) are to be recorded alongside overview and close-up photographs and crack-comparator readings. A U-Net with a pretrained ResNet34 encoder is to be fine-tuned on images from two beams and evaluated once on the third, against a classical baseline. Planning calculations to IS 456 indicate surface crack widths of roughly 0.07 to 0.23 mm over the intended load range, which led to a target close-up pixel size of about 0.03 mm. These are design values for test planning; no experimental or model results are reported in this paper.

## VIII. ENGINEERING IMPLICATIONS

Four practical points follow from the review. First, AI segmentation should be treated as a measuring aid that extracts visible crack geometry, not as a substitute for structural assessment; it cannot tell whether a crack is benign. Second, the camera set-up should be designed from the narrowest width that must be measured, with overview images used for crack pattern, length and count, and close-up images for width. Third, physical width readings remain necessary both as ground truth and as the basis for any comparison with code limits [1]; an image-derived width should be reported with its pixel size and reference. Fourth, crack metrics become most informative when they are read together with load and deflection at the same stage, because the same width can correspond to different structural states.

## IX. CONCLUSIONS

This review examined vision-based and sensor-based crack monitoring with a focus on RC beams. Learning-based segmentation outperforms classical image processing in rejecting surface noise, and encoder–decoder networks of the U-Net family provide the boundary detail needed to estimate width. Reported performance figures, however, describe different tasks, images, references and data splits, so they cannot be used to rank methods across studies. Width estimates from images are limited fundamentally by pixel size, and near code limits a one-pixel error can be a large fraction of the width. Evidence that crack images can be read alongside measured structural response comes from a small number of beam and wall studies, several of which infer load through a model or rely on a single test. The most useful next step is therefore not a new network architecture but a documented laboratory protocol that synchronises sensor readings, physical width measurements and image-derived crack metrics, specifies image resolution in advance and tests models on a withheld specimen. Such a protocol is being developed in the first author's dissertation, and its results will be reported separately.

## ACKNOWLEDGMENT

The authors thank the Department of Civil Engineering, Chhotubhai Gopalbhai Patel Institute of Technology, Uka Tarsadia University, for supporting the dissertation work from which this review was developed.

## REFERENCES

[1] Bureau of Indian Standards, *IS 456:2000 Plain and Reinforced Concrete – Code of Practice*, 4th rev. New Delhi, India: BIS, 2000.

[2] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth, "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure," *Adv. Eng. Inform.*, vol. 29, no. 2, pp. 196–210, 2015, doi: 10.1016/j.aei.2015.01.008.

[3] B. F. Spencer Jr., V. Hoskere, and Y. Narazaki, "Advances in computer vision-based civil infrastructure inspection and monitoring," *Engineering*, vol. 5, no. 2, pp. 199–222, 2019, doi: 10.1016/j.eng.2018.11.030.

[4] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Eng. J.*, vol. 57, pp. 787–798, 2018.

[5] H. S. Munawar, A. W. A. Hammad, A. Haddad, C. A. P. Soares, and S. T. Waller, "Image-based crack detection methods: A review," *Infrastructures*, vol. 6, no. 8, Art. no. 115, 2021, doi: 10.3390/infrastructures6080115.



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

[6] C. R. Farrar and K. Worden, "An introduction to structural health monitoring," *Phil. Trans. R. Soc. A*, vol. 365, no. 1851, pp. 303–315, 2007, doi: 10.1098/rsta.2006.1928.

[7] R. Hou and Y. Xia, "Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019," *J. Sound Vib.*, vol. 491, Art. no. 115741, 2021, doi: 10.1016/j.jsv.2020.115741.

[8] C.-Z. Dong and F. N. Catbas, "A review of computer vision-based structural health monitoring at local and global levels," *Struct. Health Monit.*, vol. 20, pp. 692–743, 2021, doi: 10.1177/1475921720935585.

[9] X. W. Ye, C. Z. Dong, and T. Liu, "A review of machine vision-based structural health monitoring: Methodologies and applications," *J. Sensors*, vol. 2016, Art. no. 7103039, 2016, doi: 10.1155/2016/7103039.

[10] M. Azimi, A. D. Eslamlou, and G. Pekcan, "Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review," *Sensors*, vol. 20, no. 10, Art. no. 2778, 2020, doi: 10.3390/s20102778.

[11] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang, "Automatic pixel-level crack detection and measurement using fully convolutional network," *Comput.-Aided Civ. Infrastruct. Eng.*, vol. 33, no. 12, pp. 1090–1109, 2018, doi: 10.1111/mice.12412.

[12] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, 1979, doi: 10.1109/TSMC.1979.4310076.

[13] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, 1998, pp. 130–137, doi: 10.1007/BFb0056195.

[14] S. Dorafshan, R. J. Thomas, and M. Maguire, "Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete," *Constr. Build. Mater.*, vol. 186, pp. 1031–1045, 2018, doi: 10.1016/j.conbuildmat.2018.08.011.

[15] Y.-J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Comput.-Aided Civ. Infrastruct. Eng.*, vol. 32, no. 5, pp. 361–378, 2017, doi: 10.1111/mice.12263.

[16] S. Dorafshan, R. J. Thomas, and M. Maguire, "SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks," *Data Brief*, vol. 21, pp. 1664–1668, 2018, doi: 10.1016/j.dib.2018.11.015.

[17] C. V. Dung and L. D. Anh, "Autonomous concrete crack detection using deep fully convolutional neural network," *Autom. Constr.*, vol. 99, pp. 52–58, 2019, doi: 10.1016/j.autcon.2018.11.028.

[18] Y. Liu, J. Yao, X. Lu, R. Xie, and L. Li, "DeepCrack: A deep hierarchical feature learning architecture for crack segmentation," *Neurocomputing*, vol. 338, pp. 139–153, 2019, doi: 10.1016/j.neucom.2019.01.036.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Cham, Switzerland: Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[21] D. Dais, İ. E. Bal, E. Smyrou, and V. Sarhosis, "Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning," *Autom. Constr.*, vol. 125, Art. no. 103606, 2021, doi: 10.1016/j.autcon.2021.103606.

[22] H. M. P. B. Ariyaratne, U. G. U. P. Samarasekara, E. M. R. Ekanayake, J. A. S. C. Jayasinghe, U. Jayasinghe, and A. J. Dammika, "Automated crack segmentation and width quantification using deep learning for structural health monitoring of reinforced concrete structures," *Engineer: J. Inst. Eng., Sri Lanka*, vol. 59, no. 3, pp. 1–13, 2026, doi: 10.4038/engineer.v59i3.7760.

[23] P. Iakubovskii, "Segmentation Models PyTorch," GitHub repository, 2019. [Online]. Available: https://github.com/qubvel/segmentation_models.pytorch

[24] M. R. Jahanshahi and S. F. Masri, "A new methodology for non-contact accurate crack width measurement through photogrammetry for automated structural safety evaluation," *Smart Mater. Struct.*, vol. 22, no. 3, Art. no. 035019, 2013, doi: 10.1088/0964-1726/22/3/035019.

[25] W. Qiao, H. Zhang, F. Zhu, and Q. Wu, "A crack identification method for concrete structures using improved U-Net convolutional neural networks," *Math. Probl. Eng.*, vol. 2021, Art. no. 6654996, 2021, doi: 10.1155/2021/6654996.

[26] M. Carrasco, G. Araya-Letelier, R. Velázquez, and P. Visconti, "Image-based automated width measurement of surface cracking," *Sensors*, vol. 21, no. 22, Art. no. 7534, 2021, doi: 10.3390/s21227534.

[27] R. Davoudi, G. R. Miller, and J. N. Kutz, "Data-driven vision-based inspection for reinforced concrete beams and slabs: Quantitative damage and load estimation," *Autom. Constr.*, vol. 96, pp. 292–309, 2018, doi: 10.1016/j.autcon.2018.09.024.

[28] H. Ding, C. Zhang, Y. Zhao, and J. Yu, "A load estimation method based on surface crack distribution images of reinforced concrete beams," *Buildings*, vol. 15, no. 6, Art. no. 922, 2025, doi: 10.3390/buildings15060922.

[29] M. Słoński and M. Tekieli, "2D digital image correlation and region-based convolutional neural network in monitoring and evaluation of surface cracks in concrete structural elements," *Materials*, vol. 13, no. 16, Art. no. 3527, 2020, doi: 10.3390/ma13163527.

[30] A. Rezaie, R. Achanta, M. Godio, and K. Beyer, "Comparison of crack segmentation using digital image correlation measurements and deep learning," *Constr. Build. Mater.*, vol. 261, Art. no. 120474, 2020, doi: 10.1016/j.conbuildmat.2020.120474.



(This work is licensed under a Creative Commons Attribution 4.0 International License.)