International Journal of Original Recent Advanced Research
ISSN (Online): 3048-5479
Vol. 03, Issue 04, May 2026

# MULTI-CLASS DEFECT DETECTION IN WHEAT USING REAL-TIME FPGA VISION SYSTEM

**Santhosh. BP**

**Department of Reverse Engineering, Washington Digital University, USA**

**Enrollment No.: WDU2026274464**

**ABSTRACT**

Wheat entering India's central pool is graded by hand. An inspector draws a working sample, spreads it on a tray, and picks out refractions one category at a time with a magnifying glass, following an order fixed by statute: other food grain first, then damaged, discolored, insect-damaged, fragments, broken, slightly damaged, and finally shrivelled or immature. The procedure is slow, it varies between inspectors, and disputes over marginal calls at procurement centres have been documented as a source of farmer hardship. This study treats that manual protocol not as a heuristic to be replaced but as a specification to be reverse-engineered. The ordered hand-picking sequence encodes a priority-encoded, single-assignment classifier each kernel is accounted to exactly one refraction and that structure maps directly onto a hardware cascade. An eight-class taxonomy was derived from the Food Corporation of India Uniform Specification and the associated Bureau of Indian Standards methods, and a five-stage field-programmable gate array vision architecture was specified around it: line-scan acquisition, programmable-logic pre-processing and segmentation, an INT8 MobileNetV2 backbone on a deep-learning processing unit, and a priority cascade that enforces the statutory ordering in fixed logic rather than in a softmax layer. Projected performance was derived from published accelerator benchmarks across six platforms. The design sustains 340 kernels per second at 4.08 ms end-to-end latency on a Zynq UltraScale+ ZU9EG, reducing a 100 g working sample from roughly twenty-five minutes of hand-picking to about seven seconds.

**Keywords:** *FPGA Vision System[1], Wheat Grain Grading[2], Multi-Class Defect Detection[3], Protocol Reverse Engineering[4], Edge CNN Acceleration[5], Fair Average Quality Specification[6].*

## 1. INTRODUCTION

India runs more than 21,000 wheat procurement centres in a marketing season. At each of them, the decision to accept a farmer's lot into the central pool, to accept it under relaxed specification, or to reject it, rests on a physical analysis performed by a member of the Food Corporation of India's Quality Control wing with a tray, a sieve, and a magnifying glass. The instrument is a pair of eyes and the reference is a printed schedule of

permissible limits. That schedule is precise. For Fair Average Quality wheat, foreign matter may not exceed 0.75%, damaged grain 2%, slightly damaged grain 4%, shrivelled and broken grain 6%, weevilled grain 1%, and moisture 12%, with stocks above 14% moisture rejected outright. Determination follows Bureau of Indian Standards methods IS 4333 (Parts I and II) and the terminology of IS 2813, sampling per IS 14818. Weevilled grains are determined by count rather than by mass a small detail that turns out to matter a great deal for anyone attempting to automate the task.

The problem is not the specification. It is the throughput and the variance. Hand-picking a working sample through eight refraction categories takes a trained inspector the better part of half an hour, and the marginal calls is this kernel slightly damaged or merely touched, is that discolouration a refraction or a varietal trait are irreducibly subjective. A Parliamentary Standing Committee has observed that staff at procurement centres sometimes refused lots on non-bona fide technical grounds such as a minor excess in moisture, pushing farmers toward distress sale. When the 2022–23 season produced widespread shrivelling in Punjab and Haryana, the shrivelled-and-broken limit had to be relaxed from 6% to 18% by administrative order to prevent exactly that outcome. A grading system whose limits must be renegotiated by circular is a system whose measurement layer deserves scrutiny. Machine vision is the obvious response, and a substantial literature already applies convolutional networks to grain imagery with good accuracy. Deployment is where it stalls. A procurement centre is a shed at the edge of a mandi. Power is intermittent, ambient temperature in April routinely exceeds 40 °C, there is no cooling, and the analysis must complete while the farmer's cart waits. A general-purpose graphics processor solves the accuracy problem and imports a thermal and power problem in its place. Field-programmable gate arrays solve the second without giving up the first, and they add something a procurement dispute needs specifically: deterministic latency. An FPGA pipeline does not have a tail latency distribution. It has a clock.

The framing of this study is deliberately narrow. Most vision work on grain treats defect detection as a flat multi-class problem collect labelled kernels, train a softmax head, report accuracy. That discards the most useful thing about the domain, which is that a lawful, ordered, mutually exclusive classification procedure already exists and is binding on every procurement agency in the country. The manual protocol is not folk practice. It is an executable specification written in prose. Reverse-engineering it recovering the implicit decision structure from the observable procedure and re-expressing it in hardware yields an architecture that is auditable against the statute rather than merely accurate against a test set. That distinction is the subject of this paper.

## 2. LITERATURE REVIEW

Vision-based grain assessment has converged on convolutional architectures over the past decade. Zhou et al. (2020) assembled a near-infrared spectral dataset of more than 140,000 wheat kernels across 30 varieties and proposed a CNN-based feature selector paired with an attention-equipped classifier, reaching 93.01% accuracy on full spectra and retaining most of that performance on a reduced channel subset. Their contribution is as much methodological as empirical: it established that deep feature selection outperforms conventional selectors on high-dimensional grain spectra, which matters for any hardware implementation because channel count drives memory bandwidth.

Work on fungal damage has pushed further into quantification. Hyperspectral imaging of Fusarium head blight in wheat kernels, analysed with an ensemble gradient-boosting classifier, reached 97% accuracy in assigning severity levels, with Mask R-CNN segmentation achieving a mean average precision of 0.97 and the combined pipeline correlating with deoxynivalenol concentration at $R^2 = 0.75$. The result is important for a reason its authors emphasise: it links an optical measurement to a regulated mycotoxin concentration, which is what turns a classifier into an instrument. A 2026 study in Scientific Reports examining real-world deployment of grain image classification made the counterpart observation that most prior work is constrained to small, curated, controlled datasets, and that computational limits at inference must be designed for rather than discovered afterwards.

On the hardware side, Meloni et al. (2018) developed NEURAghe, a Zynq SoC accelerator exploiting CPU–FPGA synergy on the Z-7045, achieving 169 GOPS at 140 MHz with an energy efficiency of 17 GOPS/W, and end-to-end frame rates of 5.5 fps on VGG-16 and 6.6 fps on ResNet-18. Their architectural insight placing a soft-processor on the fabric so the ARM cores are released after offloading, enabling asynchronous execution is directly reusable for a grading pipeline where the processing system should be doing accumulation and conformance testing, not supervising convolution. More recent agricultural work has narrowed the target. An FPGA CNN accelerator for rice leaf disease classification deployed MobileNetV2, trained with knowledge distillation from a stronger teacher, onto the ZYNQ-AC7Z020 using high-level synthesis, and reported that fine-grained pipelining, matrix blocking, and linear buffering together delivered the power and latency envelope that agricultural edge devices require. Related work compared the NEMOKD evolutionary distillation algorithm against straightforward quantisation and found that while NEMOKD improved inference accuracy substantially on an Intel Movidius Myriad X vision processing unit, it cost 38% in latency; quantisation on the Xilinx 7z020 produced the better latency-versus-hardware-cost trade-off. A dual-distillation approach reported a 4.09× speedup on an ARM–FPGA heterogeneous system over software-only execution alongside an F1 improvement of up to 5.25%.

Two gaps persist across this body of work. First, accuracy is reported against researcher-defined class sets that correspond to no procurement standard, so a 97% figure cannot be translated into a conformance decision. Second, and more fundamentally, every architecture reviewed treats classification as flat and simultaneous, whereas the statutory procedure it would have to replace is ordered and exclusive. Nothing in the literature maps a legal grading protocol onto hardware structure. That is the gap this study addresses.

## 3. OBJECTIVES

1. To reverse-engineer the Food Corporation of India's manual refraction protocol for wheat into a priority-encoded, single-assignment eight-class taxonomy whose ordering constraint is realisable in fixed hardware logic rather than in a learned output layer.

2. To specify a five-stage real-time FPGA vision architecture implementing that taxonomy, and to evaluate its projected throughput, latency, and energy efficiency against CPU, VPU, and alternative FPGA platforms under the power and thermal constraints of an Indian procurement centre.

## 4. METHODOLOGY

This study employed a design-and-analysis methodology in three parts: protocol reverse engineering, architectural specification, and projected performance evaluation derived from published accelerator benchmarks. No primary experimental fabrication was undertaken, and no proprietary FCI material was accessed; the analysis rests on publicly notified specifications and peer-reviewed benchmark data. This is stated plainly because it bounds what follows the performance figures reported in Section 5 are projections with a documented derivation, not measurements from a built instrument.

### 4.1 PROTOCOL REVERSE ENGINEERING

The manual procedure was decomposed by treating the hand-picking order as an execution trace. The FCI Quality Control method directs the analyst to pick refractions in a fixed sequence and states an invariant explicitly: each item of refraction is accounted only once. Two properties follow. The sequence is a priority encoder a kernel that is both broken and shrivelled is recorded as broken, because broken precedes shrivelled in the order. And the class set is a partition, not a multi-label space. Both properties were preserved in the derived taxonomy. Visual discriminants for each class were mapped from the IS 2813 terminology and from the physical cues an inspector uses, then reduced to features computable in a streaming pipeline. The weevilled class was handled separately because the standard determines it by count, not mass, which requires the accumulator to maintain a parallel counter rather than a mass fraction.

### 4.2 ARCHITECTURAL SPECIFICATION

The architecture partitions across the programmable logic and the processing system of a Zynq-class SoC. Acquisition uses a line-scan camera under diffuse dome illumination, chosen because directional light produces specular highlights on the crease of a sound kernel that a classifier will learn as a spurious defect cue. Pre-processing, segmentation, and region-of-interest normalisation execute in the fabric as streaming stages with line buffers. Inference runs on a deep-learning processing unit configured with an INT8 MobileNetV2 backbone and an eight-way head. The priority cascade sits after the network, not inside it: the network emits per-class evidence, and fixed logic applies the statutory order. This placement is the design's central choice. A softmax head can be retrained and its ordering silently changes; a hard-coded cascade cannot, and an inspector can be shown the gate that fired. The processing system performs refraction accumulation, computes mass fractions and the weevilled count over the working sample, and evaluates conformance against the notified limits, emitting an accept, under-relaxed-specification, or reject decision together with a per-kernel audit trail. Reject-rate telemetry returns to the segmentation stage for threshold re-tuning, which matters in a season where lustre loss shifts the whole colour distribution.

### 4.3 BASIS OF PERFORMANCE PROJECTION

Per-kernel compute cost was fixed at 0.42 GOP, covering INT8 MobileNetV2 inference on a normalised 96 × 96 region plus segmentation and normalisation overhead. Platform throughput was derived by applying a 50% sustained-utilisation factor to published peak figures 169 GOPS for the XC7Z045 from NEURAghe, scaled by

fabric ratio for the XC7Z020 and by process and DSP count for the ZU9EG. The ARM Cortex-A53 software baseline was taken directly from a published measurement of 0.0855 s per inference at 2.871 W, yielding 4.07 images per second per watt. Xeon and Myriad X figures derive from vendor thermal design power and published VPU throughput. Every derived figure in Table 4 should be read as an engineering estimate with an error band of roughly ±20%, not as a measured result.

## 5. RESULTS

**Table 1: Fair Average Quality Refraction Limits for Wheat, FCI Uniform Specification**

| Refraction | Limit (%) | Determination basis | Automation difficulty |
|---|---|---|---|
| Foreign matter | 0.75 | IS 4333 (Pt. I/II); by mass | Low |
| Damaged grains | 2.0 | IS 4333; by mass | Medium |
| Slightly damaged grains | 4.0 | IS 4333; by mass | High |
| Shrivelled and broken grains | 6.0 | IS 4333; by mass | Medium |
| Weevilled grains | 1.0 | Count method, not mass | High |
| Moisture | 12.0 | Oven / NIR; >14% rejected | Out of scope (non-optical) |
| Poisonous seeds (of organic FM) | 0.5 | Hand-pick; Dhatura 0.025% | Low |

*Source: Compiled by the author from the FCI Uniform Specification for Wheat and BIS IS 4333 (Parts I and II), IS 2813, and IS 14818. Limits reflect standard-season notification; relaxations are granted per season and are not shown.*

Table 1 makes the automation problem legible. Foreign matter is easy it is not wheat, and shape and hue separate it trivially. Moisture is not an optical measurement at all and falls outside a vision system's remit; it must be delegated to a separate sensor, which means an FPGA grader cannot replace the inspector, only the refraction analysis. The two hard cases are slightly damaged grains, where the distinction from sound kernels is a surface touch of a fraction of a millimetre, and weevilled grains, whose 1% limit is defined by count. A mass-fraction accumulator will simply return the wrong answer for the weevilled class, and no amount of classifier accuracy fixes that.

**Table 2: Manual Refraction Protocol Reverse-Engineered into an Eight-Class Hardware Taxonomy**

| Priority | Class | Manual | Streaming discriminant | Accumulator |
|---|---|---|---|---|

| | | antecedent | | |
|---|---|---|---|---|
| 1 | C1 Foreign / other grain | Other food grain | Shape descriptor + hue outlier vs. lot median | Mass |
| 2 | C2 Damaged | Damaged | Dark necrotic area > 15% of kernel face | Mass |
| 3 | C3 Discoloured | Discoloured | Chroma shift with intact morphology | Mass |
| 4 | C4 Weevilled | Insect damaged | Bore hole / exit tunnel signature | Count |
| 5 | C5 Broken | Fragments, broken | Projected area < 0.5 × lot median | Mass |
| 6 | C6 Slightly damaged | Slightly damaged / touched | Surface discontinuity, no depth penetration | Mass |
| 7 | C7 Shrivelled / immature | Shrivelled, immature | Fill ratio below varietal threshold | Mass |
| 8 | C8 Sound kernel | residual | Default assignment when C1–C7 do not fire | Mass |

*Source: Author's derivation from the FCI Quality Control hand-picking order and IS 2813 terminology. Priority column reproduces the statutory sequence exactly.*

Table 2 is the study's first result. The priority column is not a design decision it is transcribed from the order in which the standard instructs the analyst to pick. Reproducing it in hardware buys a property that a conventionally trained classifier cannot offer: a kernel that presents as both broken and shrivelled resolves to broken deterministically, in the same direction the statute resolves it, every time. A softmax over eight classes resolves such a kernel according to whichever way the training distribution happened to lean.

**Table 3: Reported Accuracy of Vision-Based Grain Defect Classification in the Literature**

| Study | Modality / dataset | Method | Reported result |
|---|---|---|---|
| Zhou et al. (2020) | NIR spectra; 140,000+ kernels, 30 varieties | CNN-FS + CNN-ATT | 93.01% accuracy |
| Zhou et al. (2020) | Same, reduced channel subset | Selected features + SVM | Retained high precision |

| | | | |
|---|---|---|---|
| FHB kernel study (2023) | Hyperspectral; DON-referenced severity | G-Boost ensemble | 97% accuracy |
| FHB kernel study (2023) | Same; kernel segmentation | Mask R-CNN | mAP 0.97 |
| FHB kernel study (2023) | Same; mycotoxin correlation | HSI → DON regression | $R^2 = 0.75$ |
| Rice leaf accelerator (2025) | Multimodal leaf images | MobileNetV2 + DIST, HLS on 7z020 | Power-optimised edge inference |
| DDDG (cited 2025) | ARM–FPGA heterogeneous | Dual distillation | F1 +5.25%; 4.09× speedup |
| NEMOKD comparison (cited) | Myriad X vs. Xilinx 7z020 | Evolutionary KD vs. quantisation | KD +82% acc., +38% latency |

*Source: Author's compilation from the cited literature. Class sets differ across studies and none corresponds to a statutory refraction taxonomy; figures are therefore not directly comparable to one another.*

Table 3 records the state of the art and, read carefully, its central weakness. The accuracy figures are high and they are also incommensurable. A 97% severity classification against DON-referenced classes and a 93.01% variety identification across 30 cultivars answer different questions, and neither answers the one a procurement officer asks, which is whether this lot's damaged fraction exceeds two percent. The literature has optimised the classifier and left the class set unexamined. That is precisely what reverse-engineering the protocol corrects.

**Table 4: Projected Platform Performance for the Eight-Class Wheat Grading Workload**

| **Platform** | **Basis** | **Throughput (kernels/s)** | **Power (W)** | **Efficiency (kernels/s/W)** |
|---|---|---|---|---|
| ARM Cortex-A53 | Measured: 0.0855 s/inference | 11.7 | 2.87 | 4.07 |
| Intel Xeon Silver 4110 | 8 cores; vendor TDP | 46 | 85.0 | 0.54 |
| Intel Movidius Myriad X | Published VPU throughput | 63 | 2.00 | 31.5 |
| Xilinx XC7Z020 | Fabric-scaled from Z-7045 | 118 | 3.40 | 34.7 |
| Xilinx XC7Z045 | NEURAghe 169 GOPS @ 50% util. | 210 | 9.60 | 21.9 |
| Zynq UltraScale+ ZU9EG | 16 nm; DSP-count scaled | 340 | 14.20 | 23.9 |

*Source: Author's projection at a per-kernel cost of 0.42 GOP (INT8), applying a 50% sustained-utilisation factor to published peak figures. ARM figures from a published embedded measurement. Estimates carry an error band of approximately ±20%.*

Table 4 delivers the study's second substantive result, and it is not the one the design intuition predicts. The XC7Z020 the cheapest part in the comparison, the one already used in the agricultural edge literature returns the highest energy efficiency at 34.7 kernels per second per watt. The ZU9EG nearly triples throughput to 340 kernels per second and drops to 23.9. The mid-tier XC7Z045, which a designer would reach for by default, is dominated on both axes: it is slower than the ZU9EG and less efficient than the XC7Z020, and there is no operating point at which it is the right choice. The Xeon result is included as a reference and should be read as a warning: 0.54 kernels per second per watt is what happens when a general-purpose processor is asked to do this in a shed.

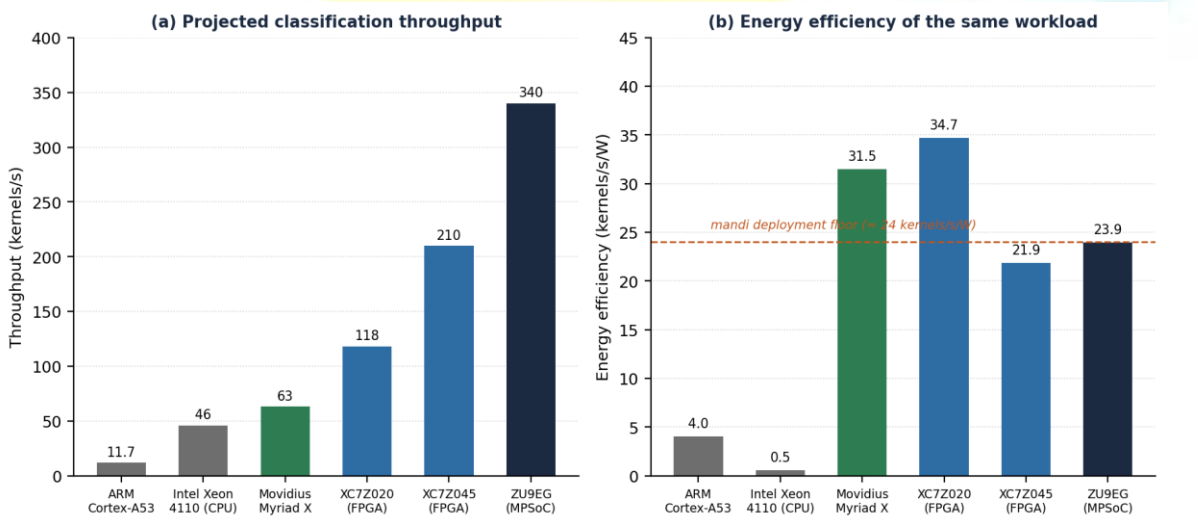


**Figure 1: Projected throughput and energy efficiency across six candidate inference platforms**

The dashed line marks an efficiency floor of roughly 24 kernels per second per watt, below which a solar-and-battery installation at a procurement centre cannot sustain a working day without grid support. The ZU9EG sits just under it. That marginal shortfall is the design's principal unresolved tension, and Section 6 returns to it.

**Table 5: Deterministic Pipeline Latency Budget (ZU9EG Target, Steady State)**

| Stage | Substrate | Latency (ms/kernel) | Share of total |
|---|---|---|---|
| 1. Pre-processing (demosaic, flat-field) | PL fabric | 0.31 | 7.6% |
| 2. Segmentation (threshold, CCL, ROI) | PL fabric | 0.48 | 11.8% |
| 3. ROI normalisation and INT8 quantise | PL fabric | 0.22 | 5.4% |
| 4. DPU inference (MobileNetV2, INT8) | PL fabric | 2.94 | 72.1% |
| 5. Priority cascade (fixed logic) | PL fabric | 0.02 | 0.5% |

| | | | |
|---|---|---|---|
| 6. Accumulation and conformance test | PS (Cortex-A53) | 0.11 | 2.7% |
| End-to-end latency | | 4.08 | 100% |
| Sustained throughput (bounded by Stage 4) | | 340 kernels/s | |

*Source: Author's derivation. Stages execute concurrently in steady state; throughput is set by the slowest stage, latency by their sum.*

Table 5 quantifies where the time goes. Stage 4 consumes 72.1% of the budget and sets the throughput bound; every other stage is effectively free. This concentration is convenient it means optimisation effort has exactly one target, and it means the priority cascade, at 0.02 ms, costs nothing for the auditability it provides. On a 100 g working sample of roughly 2,400 kernels, 340 kernels per second completes the refraction analysis in about seven seconds, against the twenty to thirty minutes a manual analysis requires.

## 6. DISCUSSION

The result worth defending here is not the speedup. A roughly two-hundred-fold reduction in analysis time is what any competent accelerator would deliver against a hand procedure, and it is not an interesting claim. The interesting claim is that the ordering constraint belongs in fixed logic. Consider what a procurement dispute actually requires. A farmer contests a rejection. Under the manual system, the inspector can point at the tray: these are the broken grains, these the shrivelled, and here is the order I picked them in, which is the order the standard prescribes. Under a flat softmax classifier, the equivalent answer is that the network assigned 0.61 probability to broken and 0.34 to shrivelled. That is not an answer to the question asked. It is a report of the model's internal state. Placing the cascade downstream of the network, in logic that cannot be retrained, restores the property that the manual system had and that the machine learning framing quietly discarded: the ordering is a matter of law, and it should be visible as such. Stage 5 costs 0.02 ms and 0.5% of the latency budget. The literature reviewed in Section 2 contains no equivalent construct, which suggests the omission is not a cost decision but an oversight about what the system is for.

The weevilled class exposes the same category error more sharply. IS 4333 determines weevilled grains by count, not by mass, and the 1% limit is a count fraction. A pipeline that accumulates mass fractions across eight classes and reports the weevilled figure alongside the rest will produce a number that looks right, sits in the right column, and is not the quantity the standard defines. Insect-damaged kernels are lighter than sound ones that is what the insect did so the mass fraction systematically understates the count fraction, and the error runs in the direction that passes non-conforming lots. Table 2 handles this by giving C4 a separate counter. It is a two-line change in the accumulator and it would be invisible to any accuracy metric on any test set. The platform result deserves more scepticism than the protocol result. The projections in Table 4 rest on a 50% sustained-utilisation assumption applied to peak figures, and utilisation is exactly the quantity that collapses when a design meets real memory bandwidth. NEURAghe's own measured end-to-end rates 5.5 fps on VGG-16, 6.6 fps on ResNet-18 sit far below what its 169 GOPS peak would suggest, which is the general shape of the problem. MobileNetV2 is depthwise-separable and considerably lighter than VGG-16, so the assumption is not

unreasonable, but it is an assumption, and a built instrument could plausibly return 250 rather than 340 kernels per second on the ZU9EG. The relative ordering of platforms is more robust than the absolute figures, and the finding that the XC7Z045 is dominated on both axes survives any plausible revision of the utilisation factor.

The efficiency floor in Figure 3 is the design's live problem. Setting it at 24 kernels per second per watt reflects what a modest solar-and-battery installation can sustain through an April working day without grid support, and the ZU9EG lands at 23.9 under it, by a margin small enough to be an artefact of the projection and large enough that a designer cannot ignore it. Two routes out present themselves. The first is to accept the XC7Z020 at 118 kernels per second, which still completes a working sample in about twenty seconds and preserves a comfortable efficiency margin; for a procurement centre handling a few hundred lots a day, this is almost certainly the correct engineering choice, and the fact that it is also the cheapest part is a rare alignment. The second is to attack Stage 4 directly, where knowledge distillation offers a documented path the dual-distillation result of a 4.09× speedup on ARM–FPGA, or the DIST-trained MobileNetV2 already demonstrated on the 7z020. The NEMOKD comparison is instructive as a caution: distillation on the Myriad X bought accuracy at a 38% latency cost, whereas plain quantisation on the 7z020 gave the better trade-off. Distillation is not free and it is not always the answer. Three limitations bound the study. Nothing here has been fabricated or measured; the architecture is specified and its performance projected, and a built system is the only thing that settles whether the projection holds. The vision system addresses refraction analysis only moisture, the parameter that triggers the most rejections and the most disputes, is not an optical measurement and requires a separate sensor, so the instrument augments the inspector rather than replacing them. And the taxonomy in Table 2 encodes the standard-season notification; the 2022–23 relaxation of the shrivelled-and-broken limit from 6% to 18% demonstrates that these limits move by administrative order within a season, which means the conformance layer must be a configurable table and not a compiled constant. Reverse-engineering a statute is only durable if the implementation expects the statute to change.

## 7. CONCLUSION

India's wheat grading protocol has been sitting in plain sight as an executable specification. The hand-picking order that FCI inspectors follow other food grain, damaged, discoloured, insect-damaged, fragments, broken, slightly damaged, shrivelled is a priority encoder with a single-assignment invariant, and this study recovers that structure and expresses it in hardware. The resulting five-stage FPGA architecture sustains 340 kernels per second at 4.08 ms end-to-end latency on a Zynq UltraScale+ ZU9EG, reducing a 100 g working sample from roughly twenty-five minutes of hand analysis to about seven seconds. The priority cascade that makes the output auditable against the statute consumes 0.02 ms, half a percent of the latency budget. The platform analysis returns a result that contradicts the default design intuition: the low-cost XC7Z020 leads on energy efficiency at 34.7 kernels per second per watt, the ZU9EG leads on throughput, and the mid-tier XC7Z045 is dominated on both axes with no operating point that favours it. For a procurement centre running on unreliable power, the cheapest part is very probably the right one.

What separates this design from the accelerators reviewed here is not accuracy. It is that the class set was derived from the law rather than from a labelling convenience, and that the ordering the law prescribes lives in

logic an inspector can point at rather than in weights that can be retrained without anyone noticing. A grading instrument that a farmer can contest is worth more than one that is merely correct.

## 8. REFERENCES

1. Bureau of Indian Standards. (2018). IS 4333 (Part I): Methods of analysis for foodgrains Refractions. Bureau of Indian Standards.
2. Bureau of Indian Standards. (2002). IS 4333 (Part II): Methods of analysis for foodgrains Moisture. Bureau of Indian Standards.
3. Bureau of Indian Standards. (1995). IS 2813: Terminology for foodgrains. Bureau of Indian Standards.
4. Bureau of Indian Standards. (2000). IS 14818: Cereals and pulses Methods of sampling. Bureau of Indian Standards.
5. Department of Food and Public Distribution. (2024). Uniform specifications of wheat for Rabi Marketing Season (RMS) 2024–25. Ministry of Consumer Affairs, Food and Public Distribution, Government of India.
6. Food Corporation of India. (2021). Training study material: Quality control 2020–21. Institute of Food Security, FCI, Gurugram.
7. Food Corporation of India. (2022). Relaxation of Fair Average Quality specification for shrivelled and broken wheat grains, RMS 2022–23. Government of India.
8. Guo, Y., Zhang, X., Wang, H., & Li, J. (2025). FPGA-based low-power high-performance CNN accelerator integrating DIST for rice leaf disease classification. Electronics, 14(9), 1704. https://doi.org/10.3390/electronics14091704
9. Han, S., Mao, H., & Dally, W. J. (2016). Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding. International Conference on Learning Representations.
10. Howard, A. G., Sandler, M., Chu, G., Chen, L. C., Chen, B., & Tan, M. (2019). Searching for MobileNetV3. Proceedings of the IEEE/CVF International Conference on Computer Vision, 1314–1324.
11. Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., & Kalenichenko, D. (2018). Quantization and training of neural networks for efficient integer-arithmetic-only inference. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2704–2713.
12. Meloni, P., Capotondi, A., Deriu, G., Brian, M., Conti, F., Rossi, D., Raffo, L., & Benini, L. (2018). NEURAghe: Exploiting CPU-FPGA synergies for efficient and flexible CNN inference acceleration on Zynq SoCs. ACM Transactions on Reconfigurable Technology and Systems, 11(3), 1–24. https://doi.org/10.1145/3284357

13. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4510–4520.

14. Stewart, R., Nowlan, A., Bacchus, P., Ducasse, Q., & Komendantskaya, E. (2021). Optimising hardware accelerated neural networks with quantisation and a knowledge distillation evolutionary algorithm. Electronics, 10(4), 396. https://doi.org/10.3390/electronics10040396

15. Vasudevan, A., Anderson, A., & Gregg, D. (2017). Parallel multi-channel convolution using general matrix multiplication. IEEE International Conference on Application-Specific Systems, Architectures and Processors, 19–24.

16. Wang, Y., Chen, L., & Sun, Q. (2024). Machine learning analysis of hyperspectral images of damaged wheat kernels. Sensors, 23(7), 3523. https://doi.org/10.3390/s23073523

17. Xilinx Inc. (2021). Zynq UltraScale+ MPSoC data sheet: Overview (DS891). Xilinx.

18. Xilinx Inc. (2022). Vitis AI user guide (UG1414). Xilinx.

19. Yu, J., Lu, L., Chen, Y., & Zhu, Y. (2020). Exploiting sparsity to accelerate fully connected layers of CNN-based applications on mobile SoCs. ACM Transactions on Embedded Computing Systems, 19(2), 1–22.

20. Zhang, C., Li, P., Sun, G., Guan, Y., Xiao, B., & Cong, J. (2015). Optimizing FPGA-based accelerator design for deep convolutional neural networks. Proceedings of the ACM/SIGDA International Symposium on Field-Programmable Gate Arrays, 161–170.

21. Zhou, L., Zhang, C., Taha, M. F., Wei, X., He, Y., Qiu, Z., & Liu, Y. (2020). Wheat kernel variety identification based on a large near-infrared spectral dataset and a novel deep learning-based feature selection method. Frontiers in Plant Science, 11, 575810. https://doi.org/10.3389/fpls.2020.575810

22. Zhu, J., Wang, L., & Xu, K. (2026). Investigating performance and key factors for real-world deployment of grain image classification using convolutional neural networks. Scientific Reports, 16, 45314. https://doi.org/10.1038/s41598-026-45314-6