BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

AI-Based Tablet and Capsule Inspection System Using YOLO, OpenCV and ESP32: A Low-Cost Visual-Screening Prototype with Embedded Alert

Suyash Sonawane[1], Raj Thakre[2], Sajid Tadvi[3], Prof. Ankita Pokale[4]

[1,2,3]Student (T.E. Undergraduate), [4]Guide
Department of Artificial Intelligence and Data Science Engineering,
Ajeenkya DY Patil School of Engineering, Pune, India

Abstract

Visual inspection of tablets and capsules is repetitive, and its outcome can vary between inspectors and over time. This paper presents a low-cost educational prototype that links a custom-trained YOLO object detector to a physical alert. A USB webcam supplies live frames to a host computer, where Python and OpenCV run Ultralytics YOLO inference on the CPU. The highest-confidence detection is mapped to a status, shown on a live display, and sent as a short ALERT or SAFE serial command to an ESP32 that drives a buzzer on GPIO 25. The detector uses weights trained on a public Roboflow Universe dataset of 11,832 images with two classes (defected and no-defect). We position the prototype against thirteen published studies on tablet, capsule and blister-pack inspection, describe its architecture, decision logic and circuit, and define an evaluation protocol. No project-specific accuracy, mAP or latency measurement is available, so none is claimed; published figures belong to the cited studies. We identify key limitations: a permissive 0.15 confidence threshold, a single-box decision rule, a SAFE fallback for unknown and no-detection cases, and a very small held-out test share (0.03%). The prototype is intended for educational demonstration and preliminary visible-defect screening only; human review remains essential.

Keywords: Tablet inspection, capsule inspection, YOLO, OpenCV, ESP32, computer vision, serial communication, pharmaceutical quality screening.

I. Introduction

A. Background

Visual screening of tablets and capsules looks for visible abnormalities such as chips, cracks, discolouration, surface damage and coating defects. Manual inspection is repetitive, and results may differ between inspectors or drift over time; sampling-based checks can also miss defects [1]–[3]. Published work has therefore applied machine vision and deep learning to tablet-surface inspection,

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

capsule classification and pharmaceutical-packaging analysis [1]–[13], and the sector is moving towards Industry 4.0 practices [14].

Most of this work relies on enclosed imaging, industrial cameras or microscopes, and its results are tied to each study's own dataset. Small student laboratories and project exhibitions need something simpler: an understandable, low-cost way to show repeatable camera-based screening of visible defects together with an immediate local alert.

B. Approach

This paper presents such a prototype. A USB webcam feeds live frames to a host computer running Python, OpenCV and a custom-trained YOLO detector [15]–[17]. The host maps the highest-confidence detection to a status and sends a short text command over a serial link to an ESP32 microcontroller, which sounds a buzzer when an item is predicted to be defective. The division of work is deliberate: the host performs the computation-intensive inference, while the ESP32 performs only a simple output action, so no model has to run on the microcontroller.

C. Contributions

- An integrated, low-cost workflow connecting YOLO-based visual prediction to a physical ESP32 alert.
- A documented description of the architecture, inference configuration, decision logic, serial protocol and circuit.
- A literature positioning against thirteen published studies and a stated research gap.
- A proposed evaluation protocol and an explicit separation of what is demonstrated from what is still unverified.

The rest of the paper is organised as follows. Section II reviews related work, Section III describes the architecture, Section IV presents the methodology, Section V defines the evaluation protocol, Section VI discusses the current status and design-level analysis, and Sections VII–IX cover responsible use, future work and conclusions.

II. Literature Review

A. Related Work

Published studies fall into a few groups. *Detection and localisation* methods use CNN or YOLO-family detectors. PharmaNet Deep [1] adds defect-guided feature fusion and an uncertainty-aware head and reports 99.4% mAP on a blister-pack dataset of 1,845 images, against 96.2% for a YOLOv8m baseline. A YOLOv8 variant with coordinate attention was applied to packaging defects in real time [4], and YOLOv5 has been used for film-coating defects, with 98.2% classification accuracy [5]. Freiermuth et al. [3] combine Faster R-CNN with deterministic OpenCV contour analysis on microscope images and keep operator review for suspicious tablets.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Classification approaches label whole images. Pathak et al. [2] report 99.7% multi-class accuracy on 25,200 augmented images of film-coated tablets with four induced defect types, while transfer-learning CNNs reach 95.71% [7] and 97.27% [8] balanced accuracy for capsules. These do not localise the defect. *Segmentation and anomaly* methods include a compact CNN that is about eight times faster than the nearest model with 3% of its parameters [6], and a patch-division convolutional autoencoder trained on normal samples [9]. Defects below the surface need other sensing: Ma et al. [10] reach 94% average accuracy for internal cracks, but from X-ray CT rather than visible-light images.

Process and review literature covers image-based coating control [11], 100% visual inspection in continuous manufacturing [12] and in-line machine-vision sensing for film coating [13]. Table I summarises the studies. Published metrics belong to each study's own data and imaging conditions and are not results of the present prototype.

TABLE I
SUMMARY OF REVIEWED STUDIES

Ref.	Focus and technique	Reported finding / limitation
[1] 2025	Blister packs and pills; CNN with feature fusion and uncertainty head	99.4% mAP (PharmaBlister); controlled enclosed imaging
[2] 2025	Film-coated tablets; CNN on 25,200 augmented images	99.7% accuracy; placebo tablets; classification only
[3] 2025	Surface defects; Faster R-CNN + OpenCV contours	No false negatives between batches; microscope setup
[4] 2024	Packaging defects; YOLOv8 with coordinate attention	Real-time YOLO use; packaging-level task
[5] 2022	Coating thickness and defects; YOLOv5	98.2% accuracy; coating-specific data
[6] 2022	Solid oral dosage forms; compact CNN segmentation	~8× faster, 3% of parameters; industrial data
[7] 2022	Defective capsules; transfer-learning CNN	95.71% balanced accuracy; no localisation
[8] 2022	Capsule quality control; CNN with explainable AI	97.27% balanced accuracy; proprietary data
[9] 2020	Pills in blisters; patch-division autoencoder	Better than other anomaly methods; normal-only training
[10] 2020	Internal tablet cracks; deep CNN on X-ray CT	94% average accuracy; needs X-ray CT
[11] 2020	Coating quality; image-based AI	Different measurement task and setup
[12] 2022	100% visual inspection of tablets	Industrial continuous-process context
[13] 2025	Review of continuous film coating	Review article; no new experiment

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

B. Research Gap and Positioning

Many reviewed systems use controlled imaging hardware [1], [3], [5], and their accuracy depends on their own dataset and defect set [1], [2], [5]. Some work adds uncertainty-aware outputs [1]. In the papers reviewed here, detection is not linked to a simple ESP32 buzzer alert. The project-specific gap addressed is therefore the integration of a trained YOLO visual-screening workflow with a low-cost embedded alert. The prototype still needs independent testing on unseen samples, robustness analysis under changing lighting and camera conditions, and a distinct uncertain/manual-review state. Table II positions the prototype against typical setups.

TABLE II

POSITIONING OF THE PROTOTYPE

Feature	Typical research / industrial setup	This prototype
Image capture	Enclosed imaging, industrial camera or microscope	USB webcam, camera index 0
Inference	CNN, YOLO, segmentation or anomaly detection	YOLO detector on host PC (CPU)
Feedback	Production automation or operator workflow	Serial command to ESP32 buzzer
Uncertainty	Some studies use uncertainty-aware methods	No separate uncertain state
Validation	Study-specific datasets and protocols	Independent results not yet available

III. System Architecture

A. Overview

The architecture has a visual path and an embedded alert path (Fig. 1). The visual path starts at the sample and webcam, passes through the host PC and the YOLO model, and yields a class and a confidence score that are shown in an OpenCV inspection window. The alert path sends a text command over USB serial to the ESP32, which controls the buzzer.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

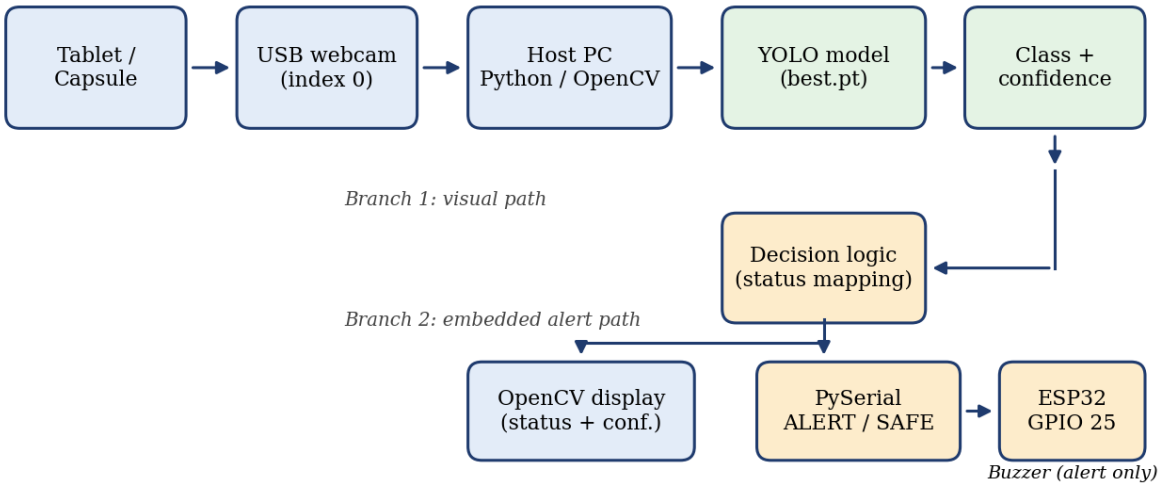


Fig. 1. Block diagram of the proposed system showing the visual path (Branch 1) and the embedded alert path (Branch 2).

B. Components

Table III lists the hardware and software. The host computer runs inference on the CPU, and no GPU is assumed. The exact ESP32 DevKit model and the buzzer type (active or passive) are still to be confirmed against the datasheets [20].

TABLE III
HARDWARE AND SOFTWARE COMPONENTS

Component	Qty.	Role / specification
USB webcam	1	Captures live frames; camera index 0
Host computer	1	Python, OpenCV and YOLO inference on CPU
ESP32 DevKit	1	Receives serial commands; drives buzzer on GPIO 25; model to be confirmed
Buzzer	1	Audible alert; (+) to GPIO 25, (–) to GND; type to be confirmed
USB cable	1	Serial link and power for the ESP32
OpenCV [18]	—	Camera capture and display
Ultralytics YOLO [17]	—	Model loading and inference
PySerial [19]	—	ALERT / SAFE commands to the ESP32
ESP32 sketch [20]	—	Reads commands, sets GPIO 25

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

C. Separation of Responsibilities

The host PC performs all image processing and inference. The ESP32 does not receive image data, store model weights or execute YOLO. This keeps the embedded task small and makes the buzzer action independent of the model implementation, provided the serial protocol is respected. The buzzer is an alert only: it does not remove, sort or physically reject an item, and no conveyor, enclosure or industrial controller is part of the prototype.

IV. Methodology

A. Dataset

The detector uses weights trained on the public Roboflow Universe dataset *Tablet Defect Detection*, Version 3, provided by FYP under the CC BY 4.0 licence. It contains 11,832 images in YOLO object-detection format with two classes, *defected* and *no-defect*. The split is 87.25% training, 12.72% validation and 0.03% testing, which corresponds to roughly 10,324, 1,505 and 3 images respectively (derived from the stated percentages). The class names expected by the code must still be confirmed from the model's own class list.

A detector can only recognise defect types represented in its training data, and rare defect types are hard to learn [3]. Near-duplicate frames or images of the same physical item must not leak across splits, because this overstates generalisation. The class distribution and annotation guidelines have not been documented here.

B. Inference Configuration

YOLO frames detection as a single-pass prediction of bounding boxes and classes [15], [16]. The prototype uses detection outputs rather than an image-classification head. Table IV lists the configured values.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

TABLE IV
INFERENCE CONFIGURATION

Parameter	Configured value / note
Weights file	runs/detect/train-2/weights/best.pt
Camera	USB webcam, index 0
Input image size	640
Confidence threshold	0.15 (permissive; may admit weak detections and raise false alarms)
Device	CPU
Inference interval	About 0.4 s (nominal, not a measured frame rate)
Detection selection	Highest-confidence box only
Serial link	115200 baud; commands ALERT / SAFE
Buzzer output	ESP32 GPIO 25

C. Decision Logic

For each inference, the box with the highest confidence is selected and its class name is mapped to a status through the model's class names. Table V gives the four states. The confidence score is a model output, not a calibrated probability that the item is defective.

TABLE V

DECISION STATES AND BUZZER BEHAVIOUR

State	Meaning	Command / buzzer
DEFECTIVE	Top box maps to a defective class	ALERT; buzzer ON
NON-DEFECTIVE	Top box maps to a non-defective class	SAFE; buzzer OFF
UNKNOWN	Class not recognised by the mapping	SAFE (fallback; limitation)
NO DETECTION	No box above the 0.15 threshold	SAFE (fallback; limitation)

The SAFE fallback for UNKNOWN and NO DETECTION is a known limitation. It can conceal uncertainty and must not be read as evidence that an item is acceptable.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

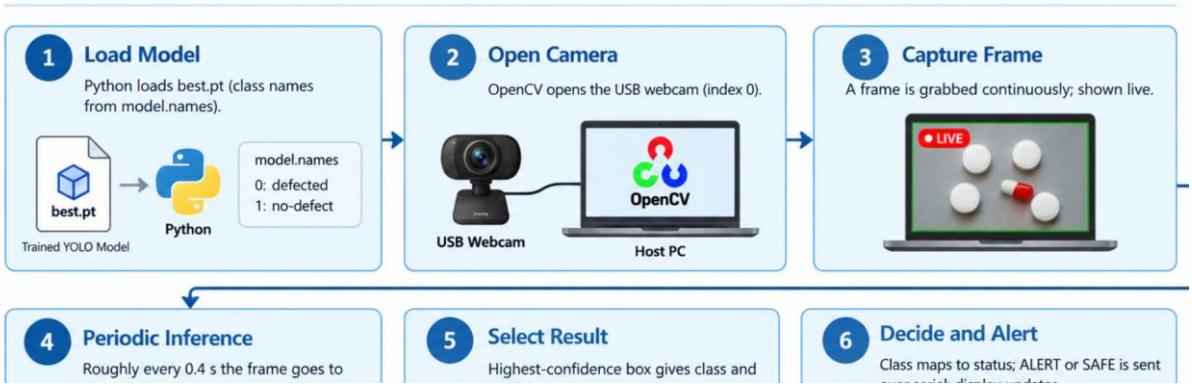


Fig. 2. Start-to-end operating cycle of the prototype: model loading, camera capture, periodic YOLO inference, result selection and ALERT/SAFE decision (schematic illustration; not a measured output).

D. Operating Cycle

Fig. 2 shows the cycle, which proceeds as follows.

1. Start the host application, load best.pt and read the class names.
2. Open camera index 0; report an error if it cannot be opened.
3. Read and display live frames continuously.
4. About every 0.4 s, run YOLO at size 640 and confidence 0.15.
5. If detections exist, select the highest-confidence box and read its class and score.
6. Map the class to DEFECTIVE or NON-DEFECTIVE; otherwise apply the fallback.
7. Send ALERT or SAFE over serial and update the display.
8. Repeat until the user exits, then release the camera and serial port.

E. ESP32 Integration and Circuit

The host sends short text commands with PySerial at 115200 baud. ALERT requests buzzer activation and SAFE requests deactivation. The ESP32 sketch listens on its serial port and sets GPIO 25 accordingly. The circuit (Fig. 3) is minimal: GPIO 25 goes to the buzzer's positive terminal, GND to its negative terminal, and the USB cable provides both power and the serial link.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

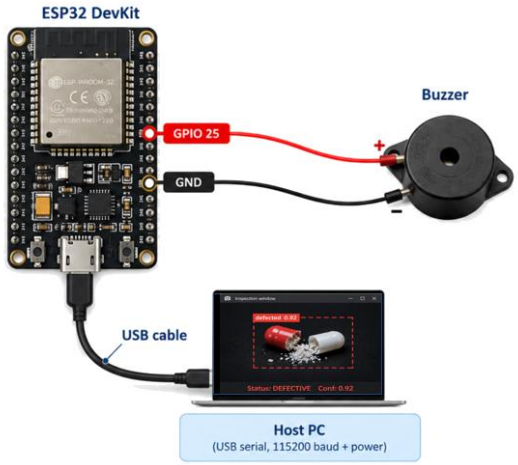


Fig. 3. Circuit and connections: ESP32 DevKit GPIO 25 and GND to the buzzer, with a USB serial link to the host PC. The laptop screen is an illustrative mock-up.

Before powering the circuit, the board's GPIO voltage and current limits and the buzzer type should be checked, and a driver transistor or series component added if the buzzer specification requires it. The command terminator and reconnect behaviour should also be confirmed from the final code, and the buzzer should be off at start-up and when the application closes.

V. Evaluation Protocol

The following protocol is defined so that the prototype can be evaluated reproducibly. It is a test plan; it does not claim that these tests have already passed.

A. Test Plan

- *Functional:* start-up (model, webcam and serial connect), live video, known defective and non-defective samples, no-detection and unknown-class behaviour.
- *Robustness:* different lighting levels, distances, orientations and backgrounds, recording changes in status and confidence.
- *Generalisation:* unseen physical tablets and capsules that were not used for training or threshold tuning.
- *Embedded:* serial command delay from host decision to ESP32 output, buzzer on/off consistency over repeated trials, and recovery after a serial disconnect.

A test passes only when the observed result matches the expected behaviour under documented conditions. Serial and buzzer tests should be repeated rather than judged from a single demonstration, and threshold changes must be tuned on validation data, not on the final test set.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

B. Metrics

Detection quality is reported per class with precision, recall and F1 score,

$$P = \mathrm{TP} / (\mathrm{TP} + \mathrm{FP}), \quad R = \mathrm{TP} / (\mathrm{TP} + \mathrm{FN}) \tag{1}$$

$$F1 = 2PR / (P + R) \tag{2}$$

where TP, FP and FN are true positives, false positives and false negatives against verified labels. A confusion matrix and per-class metrics should accompany these, together with mAP50 and mAP50–95 for object detection computed on a held-out labelled set. System-level measures are inference time and frame rate on the target computer, serial response delay, and buzzer response consistency (correct actions divided by total trials).

VI. Results and Discussion

A. Status of Results

The source material for this work describes the architecture, configuration and dataset but does not include verified project-specific accuracy, precision, recall, mAP, inference-time or serial-latency measurements, nor completed test logs. This paper therefore reports no such numbers. The figures quoted in Section II belong to the cited publications and must not be attributed to this prototype.

B. Design-Level Analysis

The architecture shows that computer vision can be coupled to a simple embedded output at low cost. Several configuration choices shape how it will behave in practice.

- *Low threshold.* A confidence threshold of 0.15 admits weak detections and can increase false alarms. It should be calibrated on validation data.
- *Single-box decision.* Only the highest-confidence box drives the status, so other items in the frame, including a defective one, may be ignored.
- *SAFE fallback.* Mapping UNKNOWN and NO DETECTION to SAFE treats absence of evidence as a pass, which is unsafe in a screening context.
- *Very small test share.* With 0.03% of the images (about three) held out, the test split cannot support reliable performance claims. An independent, larger, physically separate test set is needed.
- *Imaging conditions.* Lighting, camera angle, distance, background and sample orientation may alter predictions, and a webcam on an open desk is less controlled than the enclosed set-ups in [1], [3] and [5].

C. Threats to Validity

The dataset is public and was not collected with this camera set-up, so performance on real samples under the project's own lighting may differ. Class balance, annotation quality and duplicate frames are undocumented. The nominal 0.4 s interval is not a measured frame rate, and the exact ESP32 board and buzzer type are unconfirmed. These points must be resolved before any claim about inspection quality is made.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

VII. Responsible Use and Limitations

The prototype screens images for visible defect classes present in its training data, shows a status and confidence on screen, and sounds a buzzer for a predicted defect. It does not assess chemical composition or identity, measure potency, sterility or dissolution, prove regulatory compliance, physically reject items, or guarantee that an item predicted as NON-DEFECTIVE is safe. Its output should be described as a model prediction, not a quality decision, and a human reviewer should examine flagged and uncertain samples.

Pharmaceutical quality systems are governed by frameworks such as WHO good manufacturing practices [21], FDA process validation guidance [22], and ICH Q9(R1) and Q10 [23], [24]. These indicate the rigour expected in practice; building this student prototype does not demonstrate compliance with them. Any consequential use would require controlled imaging, documented data provenance, validated performance, risk assessment, traceability and regulatory oversight.

VIII. Future Work

- Add a distinct UNCERTAIN / MANUAL REVIEW state for unknown classes and no detections.
- Calibrate the confidence threshold and analyse false positives and false negatives on validation data.
- Evaluate on a larger independent held-out set and report class-wise metrics and a confusion matrix.
- Use controlled lighting and fixed sample positioning to reduce image variability.
- Document and expand the dataset, with tablet- and capsule-specific defect categories where data support them.
- Add inspection logging, sample identifiers and report generation, and improve serial reconnect handling.
- Consider alternative classification or segmentation models only if comparative experiments justify them, and investigate a physical reject mechanism only after separate mechanical, electrical and safety validation.

IX. Conclusion

This paper presented an educational AI-based tablet and capsule inspection workflow that combines a USB webcam, OpenCV, a custom-trained YOLO detector and an ESP32 buzzer alert. The host computer captures images, runs inference, selects the highest-confidence detection, displays a status and sends ALERT or SAFE over serial, while the ESP32 only drives the buzzer through GPIO 25. The design shows a practical separation between vision inference and embedded output control.

No independent evaluation results are available, so no project-specific accuracy or pharmaceutical-grade performance is claimed. The permissive threshold, single-box rule, SAFE fallback and very small test split are significant limitations. The next steps are to run the evaluation protocol on unseen samples, measure serial and buzzer response, add an uncertain/manual-review state, and report limitations transparently. Until then the system should be described as a preliminary visual-screening demonstration, not a validated inspection or drug-release system.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Acknowledgment

The authors thank the Department of Artificial Intelligence and Data Science Engineering, Ajeenkya DY Patil School of Engineering, Pune, for support, and Prof. Ankita Pokale for guidance.

References

[1] A. Vijayakumar, J. A. S. Koilraj and M. Rajappa, "PharmaNet Deep: Real-time pharmaceutical defect detection using defect-guided feature fusion and uncertainty-driven inspection," *Int. J. Comput. Intell. Syst.*, vol. 18, Art. no. 250, 2025, doi: 10.1007/s44196-025-00986-2.

[2] K. A. Pathak, P. Kafle and A. Vikram, "Deep learning-based defect detection in film-coated tablets using a convolutional neural network," *Int. J. Pharm.*, vol. 671, Art. no. 125220, 2025, doi: 10.1016/j.ijpharm.2025.125220.

[3] E. Freiermuth, D. Kohler, A. Hofstetter, J. Thun and M. Juhnke, "Detection and quantification of visual tablet surface defects by combining convolutional neural network-based object detection and deterministic computer vision approaches," *J. Pharm. BioTech Ind.*, vol. 2, no. 2, Art. no. 9, 2025, doi: 10.3390/jpbi2020009.

[4] A. Vijayakumar, S. Vairavasundaram, J. A. S. Koilraj, M. Rajappa, K. Kotecha and A. Kulkarni, "Real-time visual intelligence for defect detection in pharmaceutical packaging," *Sci. Rep.*, vol. 14, Art. no. 18811, 2024, doi: 10.1038/s41598-024-69701-z.

[5] M. Ficzere et al., "Real-time coating thickness measurement and defect recognition of film coated tablets with machine vision and deep learning," *Int. J. Pharm.*, vol. 623, Art. no. 121957, 2022, doi: 10.1016/j.ijpharm.2022.121957.

[6] D. Rački, D. Tomaževič and D. Skočaj, "Detection of surface defects on pharmaceutical solid oral dosage forms with convolutional neural networks," *Neural Comput. Appl.*, vol. 34, no. 1, pp. 631–650, 2022, doi: 10.1007/s00521-021-06397-6.

[7] J. Breitenbach et al., "A deep learning-based cyber-physical system for the detection of defective capsules using a transfer learning strategy," in *Proc. Pacific Asia Conf. Inf. Syst. (PACIS)*, 2022, Paper 287. [Online]. Available: https://aisel.aisnet.org/pacis2022/287

[8] D. Raab et al., "A deep learning-based model for automated quality control in the pharmaceutical industry," conference paper, 2022. [Online]. Available: https://publica.fraunhofer.de/entities/publication/3c86ca50-6a78-4667-a3f6-c0bc5a14b4c8

[9] S. Kim, Y. Jo, J. Cho, J. Song, Y. Lee and M. Lee, "Spatially variant convolutional autoencoder based on patch division for pill defect detection," *IEEE Access*, vol. 8, pp. 216781–216792, 2020, doi: 10.1109/ACCESS.2020.3041790.

[10] X. Ma et al., "Application of deep learning convolutional neural networks for internal tablet defect detection: High accuracy, throughput, and adaptability," *J. Pharm. Sci.*, vol. 109, no. 4, pp. 1547–1557, 2020, doi: 10.1016/j.xphs.2020.01.014.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

[11] C. Hirschberg, M. Edinger, E. Holmfred, J. Rantanen and J. Boetker, "Image-based artificial intelligence methods for product control of tablet coating quality," *Pharmaceutics*, vol. 12, no. 9, Art. no. 877, 2020, doi: 10.3390/pharmaceutics12090877.

[12] S. Barimani, R. Šibanc, D. Tomaževič, R. Meier and P. Kleinebudde, "100% visual inspection of tablets produced with continuous direct compression and coating," *Int. J. Pharm.*, vol. 614, Art. no. 121465, 2022, doi: 10.1016/j.ijpharm.2022.121465.

[13] D. L. Galata, O. Péterfi, M. Ficzere, B. Szabó, E. Szabó and Z. K. Nagy, "The current state-of-the-art in pharmaceutical continuous film coating – A review," *Int. J. Pharm.*, vol. 669, Art. no. 125052, 2025.

[14] N. S. Arden, A. C. Fisher, K. Tyner, L. X. Yu, S. L. Lee and M. Kopcha, "Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future," *Int. J. Pharm.*, vol. 602, Art. no. 120554, 2021, doi: 10.1016/j.ijpharm.2021.120554.

[15] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[16] A. Vijayakumar and S. Vairavasundaram, "YOLO-based object detection models: A review and its applications," *Multimed. Tools Appl.*, 2024, doi: 10.1007/s11042-024-18872-y.

[17] Ultralytics, "Object detection," *Ultralytics YOLO Documentation*. [Online]. Available: https://docs.ultralytics.com/tasks/detect/ (accessed Oct. 14, 2026).

[18] G. Bradski, "The OpenCV Library," *Dr. Dobb's J. Softw. Tools*, 2000.

[19] C. Liechti, "pySerial documentation." [Online]. Available: https://pyserial.readthedocs.io/

[20] Espressif Systems, "ESP32 series datasheet." [Online]. Available: https://www.espressif.com/en/support/documents/technical-documents

[21] World Health Organization, "WHO good manufacturing practices for pharmaceutical products: Main principles," WHO Tech. Rep. Ser., no. 986, Annex 2, 2014. [Online]. Available: https://www.who.int/publications/m/item/trs986-annex2

[22] U.S. Food and Drug Administration, "Process validation: General principles and practices," Guidance for Industry, Jan. 2011. [Online]. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/process-validation-general-principles-and-practices

[23] International Council for Harmonisation, "Quality risk management Q9(R1)," ICH Harmonised Guideline, Jan. 18, 2023. [Online]. Available: https://database.ich.org/sites/default/files/ICH_Q9%28R1%29_Guideline_Step4_2023_0126_0.pdf

[24] International Council for Harmonisation, "Pharmaceutical quality system Q10," ICH Harmonised Tripartite Guideline, Jun. 4, 2008. [Online]. Available: https://www.fda.gov/media/71553/download