

Practice and Exploration of Applying Deep Learning-based Computer Vision Technology to Microbiology Laboratory Instruction

Han Hongbo

[**Abstract**] The manual, image-intensive tasks that dominate the undergraduate microbiology laboratory—counting colony-forming units, classifying stained cells, and interpreting microscopic morphology—are subjective, time-consuming, and offer students little immediate feedback, particularly in large classes. In professional and clinical microbiology, these same tasks are increasingly performed by deep learning-based computer vision (CV) systems that now match or exceed expert accuracy on colony counting and species classification. This paper argues that the convergence of a mature enabling technology and a well-documented pedagogical problem creates an authentic opportunity for laboratory teaching reform. Adopting a review-and-framework approach rather than reporting a single controlled trial, the paper synthesizes evidence from three pieces of literature: technical CV-in-microbiology research, the pedagogy of technology-enhanced science laboratories, and the teaching-reform context of Chinese higher education. Grounded in Outcome-based Education, Bloom's revised taxonomy, and contemporary artificial-intelligence-literacy frameworks, we propose a four-rung "tool ladder" that scales from no-code image classifiers to Python-based model training on public datasets, allowing instructors to match technical demand to student readiness. We map each rung to intended learning outcomes spanning disciplinary knowledge, computational thinking, data literacy, and the critical evaluation of automated outputs. A recently published deployment at a Chinese university, in which a generative model automatically recognizes and scores students' microscopic images, is examined as a concrete precedent, and a modular course design with an aligned assessment rubric is offered for adaptation. We contend that embedding CV into microbiology instruction augments rather than replaces essential wet-laboratory skills, cultivates the interdisciplinary competencies emphasized by China's Emerging Engineering and Emerging Agricultural Science initiatives, and prepares graduates for an increasingly automated field. The paper closes by identifying limitations—including equity of access, dataset bias, and the current scarcity of controlled outcome studies—and by proposing directions for rigorous evaluation of these reforms.

[**Key words**] computer vision; deep learning; microbiology laboratory; teaching reform; Outcome-based Education; AI literacy; computational thinking; higher education

[**About the author**] Han Hongbo, affiliated with College of Biology and Chemical Engineering (College of Agriculture), Panzhuhua University.

[**Fund**] This paper is funded by "Practice and Exploration of Computer Vision Technology in Deep Learning Applied to Microbiology Experiments" (Project Number: JJ26048).

[**DOI**] <https://doi.org/10.62662/kxwxy0308001>

[**Website**] www.oacj.net

1 Introduction

The microbiology laboratory occupies a central place in the education of students across the life sciences, medicine, pharmacy, food science, and environmental engineering. It is in the teaching laboratory that abstract concepts—microbial growth, morphology, staining chemistry, and quantification—become tangible through hands-on practice. Yet the pedagogical form of the microbiology laboratory has changed remarkably little over several decades, even as the professional practice of microbiology has been transformed by automation, digital imaging, and artificial intelligence. A widening gap has opened between how microbiology is taught and how it is increasingly done.

Much of the intellectual labor of a microbiology laboratory is fundamentally visual. Students count colonies on agar plates, judge the density of growth, distinguish Gram – positive from Gram – negative cells under the microscope, and classify fungal and bacterial morphologies. These image – interpretation tasks are precisely the ones at which modern deep learning – based computer vision (CV) now excels. Since the resurgence of deep neural networks in the early 2010s, convolutional architectures have achieved expert – level or superior performance on the segmentation, detection, and classification of biological images. In microbiology specifically, deep learning models can count bacterial colonies with error rates below five percent and classify bacterial species from microscope images with accuracies exceeding ninety – nine percent.

This convergence—of a mature technology with a laboratory curriculum built around exactly the tasks that technology performs—creates an unusual opportunity for teaching reform. The opportunity is not merely to automate away tedious steps, but to use CV as an authentic, motivating vehicle for developing competencies that traditional microbiology courses rarely address: computational thinking, data literacy, and the critical evaluation of algorithmic outputs. For Chinese higher education in particular, where the Emerging Engineering Education and Emerging Agricultural Science initiatives call explicitly for interdisciplinary, innovation – oriented talent, the integration of CV into microbiology instruction aligns closely with national reform priorities.

This paper is a review and framework contribution rather than a controlled empirical study. Its purpose is to synthesize the relevant technical and pedagogical literature, to articulate a coherent theoretical rationale, and to offer instructors a practical, scalable framework for integrating deep learning – based CV into the microbiology teaching laboratory. The paper proceeds as follows. Section 2 characterizes the pedagogical problems of the traditional microbiology laboratory. Section 3 surveys the technical foundations of deep learning CV as they apply to microbiological imaging, at a level accessible to educators. Section 4 develops the pedagogical rationale, grounding the proposed reform in established learning theory. Section 5 presents the core contribution—a staged “tool ladder” and an outcome – mapped integration framework. Section 6 describes a concrete implementation, drawing on a recently published Chinese case study, and offers a modular course design. Section 7 addresses assessment and learning outcomes. Section 8 discusses limitations and ethical considerations, and Section 9 concludes.

2 Challenges in traditional microbiology laboratory instruction

To justify reform, one must first characterize the problem it addresses. Four interrelated challenges recur in the traditional microbiology teaching laboratory, and each has both a practical and a pedagogical dimension.

2.1 Subjectivity and reproducibility

Manual colony counting is the canonical example of a subjective laboratory task. When two students count the same plate, or when the same student counts on two occasions, the results frequently differ, especially where colonies overlap, cluster near the plate edge, or vary in size. The technical literature describes manual counting as “long, tedious, and subject to errors”, and this variability is not merely an inconvenience; it undermines the very notion of quantitative rigor that the exercise is meant to instill. Students learn, implicitly, that microbiological measurement is imprecise, when the deeper lesson should be that measurement error can be characterized, controlled, and reduced. Similar subjectivity attends the interpretation of Gram stains, the reading of antibiotic inhibition zones, and the classification of colony morphology.

2.2 Time, throughput, and class size

Microbiological procedures are inherently slow: cultures require overnight or multi – day incubation, and the counting or scoring of results is labor – intensive. In a large class—common in Chinese universities, where a single laboratory section may enroll dozens of students—an instructor cannot provide timely, individualized feedback on every plate or slide. The bottleneck is acute for tasks requiring expert judgment, such as verifying that a student has correctly identified a microorganism from its microscopic appearance. The consequence is that feedback arrives late, if at all, and students often proceed to the next exercise without knowing whether their previous interpretation is correct.

2.3 Limited feedback and iteration

Skilled performance in any domain depends on rapid, informative feedback that allows learners to iterate. A student learning to prepare and interpret a stained smear benefits enormously from being told, within minutes, that their staining is over – decolorized or their focus is imperfect, so that they can try again. In the traditional laboratory, the scarcity of expert attention makes such tight feedback loops impossible. Students typically prepare one smear, receive a delayed and often cursory evaluation, and move on. The opportunity for deliberate, iterative practice—the mechanism by which expertise is actually built—is largely absent.

2.4 Disconnection from contemporary professional practice

Finally, the traditional laboratory increasingly fails to reflect how microbiology is practiced. Clinical microbiology laboratories now employ automated systems for colony detection, identification, and antibiotic susceptibility testing, several of which are built on machine learning and image analysis. A graduate who has counted every plate by hand but never encountered automated image analysis is, in an important sense, underprepared. Reform that introduces CV is therefore not an optional enrichment but a step toward aligning the curriculum with the discipline's trajectory.

Crucially, deep learning CV addresses all four challenges simultaneously. It provides objective, reproducible counts and classifications; it scales without additional human labor; it can deliver near-instant feedback that enables iteration; and it introduces students to the tools of contemporary practice. This alignment between problem and solution is what makes the reform compelling.

3 Deep learning-based computer vision: technical foundations for educators

Effective integration requires that instructors understand the technology at a conceptual level, even if students engage with it through simplified tools. This section provides that grounding without assuming a computer-science background.

3.1 From classical image processing to deep learning

Before deep learning, automated colony counting and cell analysis relied on classical image processing: thresholding, edge detection, and hand-engineered features fed into conventional classifiers. These methods were brittle, requiring careful tuning for each new imaging condition. The decisive shift came with convolutional neural networks (CNNs), which learn relevant visual features directly from labeled examples rather than requiring engineers to specify them. In a direct comparison on bacterial colony counting, a CNN outperformed a carefully engineered feature-based classifier, illustrating the general advantage of learned representations. The broader significance of deep learning—its capacity to learn hierarchical representations from raw data—was articulated in the field's foundational review.

3.2 Core architectures and their microbiological applications

Four families of architecture are especially relevant to microbiology, and each maps onto a recognizable laboratory task. Image classification, in which a network assigns a whole image to a category, corresponds to identifying the species on a slide; residual networks (ResNet) are a common backbone for this task. Object detection, which locates and counts individual objects, corresponds to colony counting; the You Only Look Once (YOLO) family enables real-time detection and has been applied to plates. Semantic and instance segmentation, which delineate the precise boundaries of structures, correspond to isolating individual cells; the U-Net architecture, designed for biomedical images and able to train from very few annotated examples, is the workhorse here. Finally, specialized cell-segmentation tools such as Cellpose and, for bacterial morphologies specifically, Omnipose, provide generalist segmentation with little or no tuning.

3.3 Reported performance and public datasets

The performance now attainable is striking and is what makes the technology credible as a teaching anchor. On colony counting, a large public benchmark—the AGAR dataset, comprising eighteen thousand images and more than three hundred thousand labeled colonies—has enabled models to reach counting errors below five percent. On

species classification, the DIBaS dataset of thirty-three bacterial species has become a standard benchmark; using transfer learning from a pretrained network, one study reported average classification accuracy of ninety-nine percent. U-Net-based segmentation has been used to automate colony-forming-unit counting robustly in an industrial vaccine-production setting.

The existence of these public datasets is pedagogically important: they mean that students can train and evaluate real models on real microbiological images without needing to generate a large annotated dataset themselves. A comprehensive review of image-analysis methods for microorganism counting traces the field's evolution and catalogues these resources, and is a useful entry point for instructors.

3.4 Beyond counting: staining, morphology, and susceptibility

Colony counting and species classification are the most-cited applications, but they do not exhaust the relevance of CV to the microbiology curriculum. Gram-stain interpretation—a foundational skill that many students find difficult—has been addressed by segmentation and classification pipelines that label Gram-positive and Gram-negative cells in microscopic fields, and datasets derived from public bacterial-image collections have been used to train such models to high accuracy. This is pedagogically significant because Gram staining is exactly the kind of judgment task where students benefit from immediate, objective feedback; a model that flags an over-decolorized smear can help a student diagnose their own technique. Antibiotic susceptibility testing offers a further connection to professional practice. Automated systems that read inhibition zones or interpret growth in the presence of antimicrobials are entering clinical laboratories, and several incorporate image analysis and machine learning. Introducing students to the automated reading of a disk-diffusion plate, even in simplified form, links a classic teaching exercise to the instrumentation they will encounter in the workplace and to the pressing public-health problem of antimicrobial resistance.

Segmentation deserves particular emphasis because it underlies most quantitative image analysis. When classification asks “what is in this image?” and detection asks “where are the objects?”, segmentation asks “which pixels belong to which object?”—the precise delineation needed to measure cell size, count densely packed cells, or separate touching colonies. The U-Net architecture's ability to learn from a handful of annotated images is what makes segmentation feasible in a classroom, where students cannot annotate thousands of examples. Purpose-built tools such as Omnipose, designed for the rod-shaped and filamentous morphologies of bacteria where generalist algorithms falter, show how domain knowledge and deep learning combine. For teaching, the key insight is that a student can obtain a research-grade segmentation with a single click, freeing attention for the interpretive question of whether the result is correct.

3.5 What educators should take from this

Two points matter for teaching. First, benchmark accuracies are obtained under controlled imaging conditions; images captured on student phones under variable lighting typically perform worse, and this gap is itself a valuable teaching point about data quality and generalization. Second, the availability of no-code and low-code tools means that the conceptual power of these methods is now accessible to students who cannot yet program, while the same underlying techniques remain available for advanced students to explore in depth. This gradient of accessibility is the foundation of the framework proposed in Section 5.

4 Pedagogical rationale and theoretical frameworks

The case for integrating CV rests not only on technological opportunity but on established principles of how students learn. This section grounds the reform in three complementary frameworks.

4.1 Active and authentic learning

The strongest evidence in undergraduate science education concerns the superiority of active over passive instruction. A landmark meta-analysis of 225 studies found that active-learning approaches raised examination performance by approximately 0.47 standard deviations and that failure rates under traditional lecturing were roughly 55 percent higher than under active learning. Integrating CV, when done through student-driven image

capture, annotation, and analysis, is intrinsically active; students are not watching a demonstration but generating data, running analyses, and interpreting results. The work is also authentic in that it mirrors genuine professional practice, a quality associated with deeper engagement and transfer.

4.2 Outcome-based Education and constructive alignment

Outcome-based Education (OBE) has become the dominant curriculum-design paradigm in Chinese higher engineering and science education, reinforced by China's full membership in the Washington Accord and by national reform initiatives. OBE begins by specifying intended learning outcomes and works backward to align teaching activities and assessments with them. A CV-integrated microbiology course fits this logic naturally: outcomes such as "the student can quantify microbial growth and characterize measurement uncertainty" or "the student can critically evaluate the output of an automated classifier" are concrete, assessable, and demonstrably relevant. The OBE-CDIO model, which pairs outcome orientation with a Conceive-Design-Implement-Operate project cycle, has been applied successfully to experimental course reform and offers a ready template.

4.3 Bloom's taxonomy and AI literacy

A well-designed CV activity spans the full range of Bloom's revised cognitive taxonomy. Students remember and understand microbial categories; they apply a model to new images; they analyze why a model succeeds or fails; they evaluate whether an automated count is trustworthy; and, at the highest level, they create by training or adapting a model of their own. The evaluate and create tiers, in particular, are difficult to reach in a traditional laboratory but become natural once students are asked to judge and build automated systems.

This higher-order engagement connects to the emerging goal of artificial-intelligence literacy. Ng and colleagues conceptualize AI literacy along four dimensions: knowing and understanding AI, using and applying AI, evaluating and creating with AI, and engaging with AI ethics. A CV-integrated microbiology course can advance all four, giving students not a generic exposure to AI but a domain-embedded, practice-based literacy that is more likely to transfer. Because AI is reshaping the students' future profession, cultivating the capacity to use these tools critically—rather than either rejecting them or trusting them blindly—is itself an important learning outcome.

4.4 The Chinese teaching-reform context: cultivating interdisciplinary talent

The reform proposed here does not arise in a vacuum; it resonates strongly with the trajectory of Chinese higher education. Over the past two decades, China has undertaken one of the most ambitious programs of engineering- and science-education reform in the world, introducing outcome-based curricula, project-based learning, and the CDIO model, and achieving full membership in the Washington Accord. The Emerging Engineering Education initiative (新工科) explicitly calls for interdisciplinary, innovation-oriented programs that fuse traditional disciplines with information technology and artificial intelligence, and the parallel Emerging Agricultural Science initiative (新农科) extends this logic to the agricultural and life sciences. A microbiology course that teaches students to apply computer vision sits precisely at the intersection these initiatives seek to cultivate: it joins a classical experimental science to modern computational methods, producing graduates who are neither pure biologists nor pure programmers but competent at the interface.

This interface is where much of the future work of the life sciences will occur. Microbiology is becoming a data-rich discipline, and the graduates best positioned to contribute are those who can move fluently between the wet laboratory and the computer. Chinese universities have begun to act on this recognition: the practice of teaching reform in microbiology, including outcome-based redesign and the integration of digital and blended methods, is an active area documented in the domestic literature, and the deployment of an AI image-scoring model in an undergraduate microbiology laboratory discussed later in this paper is a concrete instance. Situating the present framework within this reform ecosystem is important for two reasons. First, it means that instructors adopting the framework can align it with existing institutional priorities and funding streams rather than justifying it from scratch. Second, it means that the competencies the framework develops—computational thinking, data literacy, and critical engagement with AI—are not peripheral enrichment but central to the kind of talent the national reform agenda is

designed to produce.

5 A framework for integration: the tool ladder

The central practical contribution of this paper is a staged framework that matches technical demand to student readiness, allowing the same conceptual content to be delivered to novices and advanced students alike. We describe it as a four-rung “tool ladder”. Instructors may enter at any rung and ascend as far as their students’ preparation and course goals permit.

5.1 Rung one: no-code classification

The first rung uses browser-based, no-code tools—most notably Google’s Teachable Machine—that let students train an image classifier by uploading example images into labeled categories. Within minutes, and without writing any code, students can build a classifier that distinguishes, say, yeast from bacterial cells, or Gram-positive from Gram-negative morphologies, from their own microscope photographs. The pedagogical value lies in making abstract concepts concrete: students directly experience what “training data”, “classes”, “inference”, and “overfitting” mean when their model misclassifies an image because their training set was too small or unrepresentative. This rung is appropriate for first-year students and for courses where the primary discipline is biology rather than computation.

5.2 Rung two: guided quantification with ImageJ/Fiji

The second rung introduces ImageJ and its distribution Fiji, the standard open-source platform for biological image analysis. Students use ImageJ to count colonies, measure inhibition zones, or quantify stained area in a guided, semi-automated workflow, and can progress to writing simple macros that automate repetitive steps. Ready-made, citable teaching materials exist for exactly this purpose. This rung develops quantitative reasoning and a first taste of programmatic automation while remaining within a tool built for biologists.

5.3 Rung three: reproducible pipelines with CellProfiler

The third rung uses CellProfiler, which allows students to assemble modular, reproducible image-analysis pipelines without programming and to batch-process many images at once. Here the learning objective shifts toward reproducibility and scale: students confront the difference between analyzing one image by hand and constructing a pipeline that will process hundreds identically. This rung introduces the professional value of documented, repeatable analysis workflows.

5.4 Rung four: model training in Python

The fourth and highest rung has students work directly in Python, using libraries such as Open CV, TensorFlow, or PyTorch, together with annotation platforms and the public DIBaS and AGAR datasets, to train and evaluate their own deep learning models. A capstone project might task students with training a colony detector and reporting its counting error, or fine-tuning a pretrained network to classify bacterial species and analyzing its confusion matrix. This rung is suited to advanced undergraduates, to interdisciplinary programs, and to the Emerging Engineering and Emerging Agricultural Science tracks that explicitly seek to fuse biological and computational competencies. It develops genuine, transferable skill in machine learning practice.

5.5 Scaffolding, prerequisites, and instructor readiness

A tool ladder is useful only if students and instructors can climb it. Two practical concerns govern successful implementation. The first is scaffolding: each rung should build on the vocabulary and intuitions established by the one below. A student who has experienced, at rung one, why a classifier trained on ten images performs poorly is far better prepared to appreciate, at rung four, why a model needs a large and representative training set. Instructors should resist the temptation to leap directly to Python model training with students who have not first grasped the underlying concepts through no-code exploration; the conceptual foundation laid at the lower rungs is what makes the upper rung comprehensible rather than merely procedural.

The second concern is instructor readiness. Microbiology instructors are not, in general, machine-learning specialists, and a reform that presumes deep technical expertise on the part of every teacher will not scale. The tool

ladder is designed to mitigate this; the lower rungs require no programming and little more than the willingness to learn a browser-based tool, while the upper rung can be co-taught with colleagues from computer science or delivered in interdisciplinary programs where computational expertise is already present. Institutions can support adoption by developing shared teaching materials, providing professional development, and cultivating communities of practice in which instructors exchange datasets, exercises, and lessons learned. The existence of citable, ready-made teaching resources for image analysis lowers this barrier considerably.

5.6 Mapping rungs to outcomes

The rungs are not merely a difficulty gradient; each foregrounds different intended outcomes. The lower rungs emphasize conceptual understanding of what machine vision is and disciplinary reinforcement of microbiology. The middle rungs emphasize quantitative reasoning and reproducibility. The upper rung emphasizes computational thinking, model evaluation, and the capacity to create. Across all rungs, a common thread is the critical evaluation of automated outputs—the disposition to ask whether a machine’s count or classification should be trusted, and to detect its failure modes. This disposition, aligned with the highest tiers of Bloom’s taxonomy and with the evaluative dimension of AI literacy, is the single most important competency the framework aims to cultivate.

6 Implementation: a case study and a modular course design

6.1 A documented precedent

The framework is not hypothetical. At Shanghai Jiao Tong University, researchers in the School of Life Sciences and Biotechnology developed a generative artificial-intelligence model, informally named “Identify Microbe” (识菌), and deployed it in an undergraduate microbiology teaching laboratory. The model automatically recognizes and scores the microscopic smear and staining images that students submit—covering bacteria, molds, actinomycetes, and yeast—returning targeted feedback within seconds. In doing so it directly resolves the feedback and throughput bottlenecks identified in Section 2; rather than waiting for an instructor to inspect each slide, students receive immediate, objective evaluation and can iterate through multiple staining attempts in a single session. The deployment was supported by the university’s “AI + Education” initiative, situating it within an institutional reform agenda rather than an isolated experiment. A companion effort in the same program applied generative AI to project-based teaching of a Python course for biology majors, illustrating how CV and programming instruction can reinforce one another.

This case is instructive precisely because it operationalizes the pedagogical logic advanced here; it uses CV not to replace the wet-laboratory act of preparing a smear, but to intensify the feedback loop around it, thereby enabling the deliberate, iterative practice that traditional laboratories cannot support.

6.2 A modular course design

Building on this precedent and the tool ladder, we outline a modular design that instructors can adapt to their students’ level and their institution’s resources. The design assumes a standard microbiology laboratory course into which CV modules are inserted, rather than a wholesale replacement.

Module A—Orientation and motivation. A short introductory session frames the problem; students count the same plate manually, compare their counts, and confront the variability directly. They are then shown an automated count, motivating the reform from their own experience of subjectivity.

Module B—No-code classification. Using Teachable Machine, students photograph their own stained slides and train a classifier, learning the vocabulary of machine learning and confronting the consequences of poor training data. This module targets conceptual understanding and requires no programming.

Module C—Quantification. Using ImageJ/Fiji, students automate colony counting or zone measurement and compare automated with manual results, quantifying the agreement and discussing sources of error. This module targets quantitative reasoning.

Module D—Capstone (optional, advanced). Students working at the upper rung train or fine-tune a model on the DIBaS or AGAR dataset, evaluate its performance, and present a critical analysis of where and why it fails.

This module targets computational thinking and model evaluation, and is well suited to interdisciplinary cohorts.

Throughout, the wet-laboratory core—aseptic technique, culturing, staining, and microscopy—remains intact. The CV modules attach to it, deepening rather than displacing the hands-on experience. Institutions with limited computational infrastructure can implement Modules A through C using only a web browser and free software, while those with greater resources can extend to Module D.

6.3 Practical logistics and sequencing

Several logistical choices determine whether the modules succeed in practice. Image acquisition is the first; students need a consistent way to photograph plates and slides, whether through smartphone adapters on microscope eyepieces, flatbed scanners for plates, or dedicated imaging stations. Standardizing acquisition improves model performance and, more importantly, teaches students that reproducible data collection is the foundation of reproducible analysis. The variability introduced by inconsistent lighting or focus is not merely a nuisance to be eliminated but a phenomenon to be studied, illustrating why professional laboratories invest in controlled imaging. Data management is the second concern: even a modest class generates hundreds of images, and establishing clear conventions for naming, labeling, and storing them models good scientific data practice.

Sequencing across a semester matters as well. A workable pattern introduces the orientation module early, so that the motivating experience of manual variability is fresh; interleaves the no-code and quantification modules with the corresponding wet-laboratory exercises, so that students analyze their own freshly generated images rather than stock photographs; and reserves the capstone for the latter part of the course, when students have both the disciplinary knowledge and the conceptual grounding to attempt it. Aligning each CV module with the wet-laboratory exercise it analyzes is what makes the integration feel coherent rather than bolted on; students count the very colonies they cultured and classify the very smears they stained, closing the loop between hands-on practice and computational analysis.

7 Assessment and learning outcomes

Constructive alignment demands that assessment measure the outcomes the reform is meant to produce, not merely disciplinary recall. We propose an assessment approach organized around four outcome clusters and note the kinds of evidence that support expected gains.

7.1 Outcome clusters and rubrics

The first cluster, disciplinary knowledge, is assessed conventionally through students' correct identification and quantification of microorganisms. The second, quantitative and data literacy, is assessed through students' ability to compare automated and manual results, report agreement, and reason about measurement error. The third, computational thinking, applies chiefly at the upper rung and is assessed through the design, training, and documentation of an analysis pipeline or model. The fourth, and most distinctive, is the critical evaluation of automated outputs: students are explicitly credited for detecting and explaining a model's errors—an over-count where colonies merge, a misclassification driven by staining artifact, or a failure to generalize to a new imaging condition. A rubric that rewards the identification of AI failure modes operationalizes the highest tier of Bloom's taxonomy and the evaluative dimension of AI literacy, and it guards against the uncritical acceptance of machine output that is itself a risk of automation.

7.2 Course-based research and engagement

Where resources permit, CV modules can be embedded within a course-based undergraduate research experience (CURE), in which students investigate a genuine, open question rather than confirming a known result. CUREs in microbiology have been shown to produce gains in self-efficacy, conceptual knowledge, and persistence in science. Framing a CV project as authentic inquiry—can an automated classifier reliably distinguish these environmental isolates?—couples the engagement benefits of research with the skill development of computational analysis.

7.3 Expected outcomes and the evidence base

Because controlled studies of CV integration specifically are still scarce, expectations must be inferred from adjacent evidence. The active-learning meta-analysis predicts improved performance and reduced failure where instruction becomes more active. Meta-analytic evidence on virtual and computational laboratories reports moderate-to-large gains, with the largest effects on learning motivation and engagement, while cautioning that such tools complement rather than replace hands-on work. Studies of AI teaching assistants in undergraduate biology find that students value automated feedback but continue to prefer human guidance for nuanced judgment, underscoring that CV should augment, not supplant, the instructor. Taken together, these bodies of evidence support cautious optimism while highlighting the need for direct, controlled evaluation of the specific reforms proposed here.

7.4 Transferable skills and career relevance

A final dimension of assessment concerns the durability and transferability of what students learn. The specific tools taught—a particular no-code platform, a particular library—will change; the underlying competencies should not. A student who has grasped the logic of training data, model evaluation, and the critical scrutiny of automated output possesses a mental model that transfers to any future encounter with machine learning, in microbiology or beyond. This is why the framework foregrounds concepts over tools and evaluation over mere operation. When assessment credits students for reasoning about why a model behaves as it does, rather than only for producing a correct output, it measures the transferable understanding that will still matter after today's software is obsolete. In career terms, graduates who can collaborate with computational systems—interpreting, validating, and questioning their output—are precisely those the automated microbiology laboratory of the near future will need, whether in clinical diagnostics, food safety, environmental monitoring, or research. Framing the reform's outcomes in these durable terms also strengthens the case for its adoption to students, who are more motivated when they perceive the direct relevance of a skill to their professional futures.

8 Challenges, limitations, and ethical considerations

Responsible advocacy for reform requires candor about its limits. Several considerations temper the enthusiasm of the preceding sections.

8.1 Complementarity, not replacement

The most important caveat is that CV augments but cannot replace the wet-laboratory core of microbiology education. The tactile skills of aseptic technique, streaking for isolation, staining, and microscopy remain irreducible, and the meta-analytic literature on virtual laboratories is explicit that computational tools currently cannot fully substitute for hands-on experience. The framework proposed here is deliberately additive, attaching CV modules to an intact practical curriculum.

8.2 Equity and access

No-code and browser-based tools substantially lower the barrier to entry, but disparities remain. The upper rung of the tool ladder assumes access to adequate computing and, ideally, graphics-processing hardware, which is unevenly distributed across institutions. Instructors should design courses so that essential outcomes are attainable at the lower rungs, treating the computational capstone as an extension rather than a requirement, to avoid disadvantaging students at less-resourced institutions.

8.3 Dataset bias and generalization

Models are only as good as their training data. Public datasets, and models trained on them, may not generalize to the imaging conditions, microorganisms, or staining protocols of a particular laboratory, and much clinical-microbiology AI has been developed on data from a narrow range of settings. Rather than hiding this limitation, the framework turns it into a learning objective: confronting students with a model's failure to generalize is among the most effective ways to teach data quality, bias, and the boundaries of automation.

8.4 Automation bias and ethics

A subtler risk is that students, presented with a confident automated output, cease to think critically—the

phenomenon of automation bias. The assessment design proposed in Section 7, which rewards the detection of model errors, is a direct countermeasure. More broadly, engaging students with the ethical dimensions of AI—including questions of accountability when an automated system errs in a clinical context—fulfills the ethics dimension of AI literacy and prepares students for the responsible use of these tools in practice.

8.5 The evidence gap

Finally, and candidly, rigorous, controlled studies that measure learning outcomes from integrating deep learning CV specifically—as distinct from virtual laboratories or generic active learning—remain scarce. The documented Chinese deployment is a descriptive precedent rather than a controlled trial. This paper therefore frames its contribution as a practice-and-exploration framework that also issues a call: the reforms it proposes should be evaluated with pre-post designs, comparison groups, and validated instruments, so that the field can move from plausible rationale to demonstrated effect.

9 Conclusion and future directions

The undergraduate microbiology laboratory is built around visual, interpretive tasks—counting, classifying, and judging—that deep learning-based computer vision now performs with expert-level accuracy. This convergence of a mature technology with a curriculum organized around exactly the tasks it addresses creates a rare and compelling opportunity for teaching reform. This paper has argued that the opportunity should be seized not primarily to automate tedium, but to cultivate competencies that traditional courses neglect: quantitative and data literacy, computational thinking, and, above all, the critical evaluation of automated outputs.

The framework offered here—a four-rung tool ladder scaling from no-code classifiers to Python-based model training, mapped onto intended outcomes and grounded in Outcome-based Education, Bloom's taxonomy, and AI-literacy theory—is designed to be adaptable across institutions and student levels. A documented deployment at a Chinese university demonstrates that the approach is not merely aspirational, and the modular course design shows how CV modules can attach to an intact wet-laboratory curriculum, augmenting rather than replacing it. Such reform aligns squarely with the Emerging Engineering and Emerging Agricultural Science initiatives that call for interdisciplinary, innovation-oriented talent, and it narrows the widening gap between how microbiology is taught and how it is practiced.

Two directions for future work stand out. The first is rigorous evaluation: the field needs controlled studies, using validated instruments and comparison groups, to establish the effect of CV integration on disciplinary learning, computational competence, and AI literacy. The second is the development and sharing of open pedagogical resources—curated datasets, annotated teaching materials, and ready-to-use modules—that lower the barrier for instructors and promote equity across institutions. If these are pursued, the integration of computer vision into microbiology instruction can mature from a promising exploration into an evidence-based standard, preparing a generation of microbiologists to work fluently and critically alongside the intelligent tools that are reshaping their discipline.

References:

- [1] Beznik, T. , Smyth, P. , de Lannoy, G. , Lee, J. A. Deep Learning to Detect Bacterial Colonies for the Production of Vaccines[J]. *Neurocomputing*, 2022, 470(1) : 427–431.
- [2] Carpenter, A. E. , Jones, T. R. , Lamprecht, M. R. , Clarke, C. , Kang, I. H. , Friman, O. , Sabatini, D. M. CellProfiler: Image Analysis Software for Identifying and Quantifying Cell Phenotypes [J]. *Genome Biology*, 2006, 7(10).
- [3] Cutler, K. J. , Stringer, C. , Lo, T. W. , Rappez, L. , Stroustrup, N. , Brook Peterson, S. , Mougous, J. D. Omnipose: A High-precision Morphology-independent Solution for Bacterial Cell Segmentation[J]. *Nature Methods*, 2022, 19(11) : 1438–1448.
- [4] Ding, T. , Tang, Y. W. , Hao, X. Application of Artificial Intelligence in Clinical Microbiology: From

Research to Practice[J]. *iLABMED*, 2025, 3(4): 405–416.

[5] Dong, S. , Guo, Y. , Yang, T. Generative AI-assisted Project-based Teaching of a Python Course for Biology Majors[J]. *Journal of Biology*, 2025, 42(4): 27–30.

[6] Ferrari, A. , Lombardi, S. , Signoroni, A. Bacterial Colony Counting with Convolutional Neural Networks in Digital Microbiology Imaging[J]. *Pattern Recognition*, 2017, 61: 629–640.

[7] Freeman, S. , Eddy, S. L. , McDonough, M. , Smith, M. K. , Okoroafor, N. , Jordt, H. , Wenderoth, M. P. Active Learning Increases Student Performance in Science, Engineering, and Mathematics[J]. *Proceedings of the National Academy of Sciences*, 2014, 111(23): 8410–8415.

[8] He, K. , Zhang, X. , Ren, S. , Sun, J. Deep Residual Learning for Image Recognition[A]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [C]. 2016: 770–778.

[9] LeCun, Y. , Bengio, Y. , Hinton, G. Deep Learning[J]. *Nature*, 2015, 521: 436–444.

[10] Majchrowska, S. , Pawłowski, J. , Guła, G. , Bonus, T. , Hanas, A. , Loch, A. , et al. AGAR: A Microbial Colony Dataset for Deep Learning Detection[EB/OL]. <https://arxiv.org/abs/2108.01234>, 2021–8–3.

[11] Ng, D. T. K. , Leung, J. K. L. , Chu, S. K. W. , Qiao, M. S. Conceptualizing AI Literacy: An Exploratory Review[J]. *Computers and Education: Artificial Intelligence*, 2021, 2.

[12] Redmon, J. , Divvala, S. , Girshick, R. , Farhadi, A. You Only Look Once: Unified, Real-time Object Detection[A]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [C]. 2016: 779–788.

[13] Ronneberger, O. , Fischer, P. , Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation[A]. *Proceedings of the Medical Image Computing and Computer-assisted Intervention (MICCAI 2015)* [C]. 2015: 234–241.

[14] Schindelin, J. , Arganda-Carreras, I. , Frise, E. , Kaynig, V. , Longair, M. , Pietzsch, T. , Cardona, A. Fiji: An Open-source Platform for Biological-image Analysis[J]. *Nature Methods*, 2012, 9(7): 676–682.

[15] Schneider, C. A. , Rasband, W. S. , Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis[J]. *Nature Methods*, 2012, 9(7): 671–675.

[16] Stirling, D. R. , Swain-Bowden, M. J. , Lucas, A. M. , Carpenter, A. E. , Cimini, B. A. , Goodman, A. CellProfiler 4: Improvements in Speed, Utility and Usability[J]. *BMC Bioinformatics*, 2021, 22(1).

[17] Stringer, C. , Wang, T. , Michaelos, M. , Pachitariu, M. Cellpose: A Generalist Algorithm for Cellular Segmentation[J]. *Nature Methods*, 2021, 18(1): 100–106.

[18] Talo, M. An Automated Deep Learning Approach for Bacterial Image Classification[EB/OL]. <https://arxiv.org/abs/1912.08765>, 2019–12–4.

[19] Tischer, C. , Halavatyi, A. Bioimage Analysis Teaching Material for ImageJ[EB/OL]. <https://doi.org/10.5281/zenodo.2633701>, 2019–4–9.

[20] Wu, Z. The Giant Leap in the Development of Higher Engineering Education in China[J]. *ECNU Review of Education*, 2025, 8(2).

[21] Zhang, J. , Li, C. , Rahaman, M. M. , Yao, Y. , Ma, P. , Zhang, J. , Grzegorzec, M. A Comprehensive Review of Image Analysis Methods for Microorganism Counting: From Classical Image Processing to Deep Learning Approaches[J]. *Artificial Intelligence Review*, 2022, 55(4): 2875–2944.

[22] Zhang, X. , Li, B. Reinventing Microbiological Experimental Teaching with the AI “Identify Microbe” Model[J]. *Journal of Biology*, 2025, 42(4): 9–14.

[23] Zieliński, B. , Plichta, A. , Misztal, K. , Spurek, P. , Brzywczy-Włoch, M. , Ochońska, D. Deep Learning Approach to Bacterial Colony Classification[J]. *PLOS ONE*, 2017, 12(9).