Educational and Humanities
ISSN Print: 3104-4352
ISSN Online: 3104-4360
CD ISSN: 3104-4379
https://doi.org/10.63313/EH.9120

Exploring Reform in a Progressive, Project-Driven Deep Learning Practical Course: A Case Study of an Intelligent Image Classification System

Tong Su. Siyuan Bei*
School of Computer Science and Artificial Intelligence, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China
Email: 20250029@lixin.edu.cn

How to cite this paper: Su, T., & Bei, S. (2026). Exploring reform in a progressive, project-driven deep learning practical course: A case study of an intelligent image classification system. Educational and Humanities, 5(1), 40–49. ISSN Print: 3104-4352, ISSN Online: 3104-4360.
https://doi.org/10.63313/EH.9120
Published: 2026-10-26

Copyright © 2026 by author(s) and Erytis Publishing Limited.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
CC BY 4.0

Abstract

Addressing issues such as fragmented laboratory sessions, a disconnect between theory and practice, and one-dimensional assessment methods in deep learning practical courses, this paper explores a progressive, project-driven teaching reform centered on the iterative development of an "intelligent image classification system." The course restructured scattered experiments into four progressive projects focused on modeling from scratch, transfer learning, few-shot fine-tuning, and competition practice. By aligning with the competency requirements of the Lanqiao Cup competition, it promoted the integration of competition and coursework. The implementation incorporated AI programming assistants and large language models, while the evaluation established a five-dimensional formative assessment system covering the project process, use of AI tools, technical competence, final defense, and personal growth portfolios. Pilot results show that all four projects met the predetermined requirements. Students' engineering practice skills and technical decision-making abilities have significantly improved, and the competition pass rate exceeded historical levels, providing a valuable reference for reforming similar practical courses.

Keywords

Project-Based Learning; Deep Learning; Progressive Projects; Image Classification; Integration of Competitions and Courses

1. Introduction

Deep learning practical courses are a core practical component of majors such as Intelligent Science and Technology and Computer Science and Technology, characterized by complex theoretical frameworks, high practical demands, and rapid technological iteration. The "Outline of the Plan for Building an Education Powerhouse (2024–2035)" proposes "promoting the use of artificial intelligence to drive educational transformation," setting higher standards for strengthening

practical education tailored to industry needs. However, in actual teaching practice, such courses still commonly suffer from issues such as fragmented experiments, a disconnect between classroom training and industry and competitions, and an evaluation system focused primarily on summative outcomes—making it difficult to reflect the process of students' competency development.

Resolving these issues requires a systematic restructuring of the curriculum. Specifically, the theoretical basis for progressive project design—which addresses the problem of fragmented experiments—lies in the following: the Zone of Proximal Development theory requires that task difficulty be gradually increased, while cognitive load theory suggests breaking down complex tasks into progressively more difficult subtasks to avoid cognitive overload. Centered on Project-Based Learning (PBL), various explorations have emerged in China within deep learning and related courses, including cross-disciplinary projects to enhance practical skills, blended online and offline implementation, and research-based teaching combined with a tiered instructional system. However, these efforts often treat projects as isolated entities, lacking systematic discussion of the progressive relationship between skills across projects. Regarding the integration of teaching, practice, and competition, tiered experimental instruction combining "teaching, practice, and competition" has been explored; however, the alignment of competition skills with classroom projects often remains at a macro level. In terms of artificial intelligence tool applications, large language models have been introduced into programming courses, but practical solutions are still lacking regarding how AI tools can be integrated with formative assessment and how to prevent integrity risks such as "direct copy-and-paste".

To address these research gaps, this paper takes the iterative development of an "intelligent image classification system" as its central thread to reconstruct a four-stage progressive project chain. It integrates AI programming assistants and large language models into the curriculum, establishes a comprehensive evaluation system to support the entire process, and conducts exploratory practice in a single pilot class. This paper focuses on answering three questions: (1) Can the four-stage progressive project chain support students' progression from basic modeling to comprehensive practical application? (2) Can the alignment between classroom projects and competitions translate into observable competition performance? (3) Can AI tool assistance and end-to-end assessment enhance students' engineering practice capabilities and their ability to make informed technical decisions?

2. Curriculum Reform Design

2.1. Teaching Challenges

The course's original teaching resources already covered six major modules—including CIFAR-10 image classification, the use of pre-trained models, transfer learning fine-tuning, style generation, and Kaggle competitions—providing

relatively comprehensive theoretical coverage. However, three key issues emerged during the teaching process. First, the experiments are relatively isolated; each module is designed as a standalone unit based on specific knowledge points. Students can complete the tasks by following the lab manuals step by step, but they lack in-depth reflection on engineering decision-making questions such as "Why choose this model?" and "How to scientifically evaluate its merits and demerits," making it difficult to form a complete cognitive chain from problem definition to solution iteration. Second, classroom training is disconnected from real-world industry scenarios and academic competitions. While students master theories such as convolutional neural networks and transfer learning, they generally struggle with issues such as blind model selection, difficulties in training and debugging, and an inability to make reasonable engineering trade-offs under constraints of limited computing power and data. Third, evaluation focuses primarily on the final code execution results or lab reports, lacking tracking of process-oriented performance—such as problem decomposition, collaboration, and the proper use of AI tools—making it difficult to reflect the progression of competency development.

2.2. Reform Approach

To address these pain points, this course restructures its practical teaching content around the iterative development of an intelligent image classification system. To resolve the issue of fragmented experiments, the original six modules have been reorganized into four projects with progressively increasing difficulty and skill requirements, replacing isolated, verification-based experiments with a progression of difficulty and skills across projects; To address the disconnect between classroom instruction and competitions, the course references the competency requirements of the Lanqiao Cup AI Competition—including data processing, model building, optimization techniques, and result presentation—during the project design phase. By embedding these competition-specific competencies into the driving questions and evaluation criteria of each project, the course ensures alignment between classroom training and real-world competition scenarios; To address the issue of one-dimensional assessment, a five-dimensional formative assessment system was established, covering the project process, use of AI tools, technical skills, final defense, and personal growth portfolios. The correspondence between the three categories of countermeasures and their implementation vehicles is summarized in Table 1.

Table 1. Course Teaching Pain Points and Reform Strategies

Teaching Challenges	Reform Strategies	Implementation Vehicles
Fragmented experiments and lack of a systems engineering perspective	Progressive project chain design	Four-Stage Progressive Projects (Projects 1 through 4)
Disconnect from Industrial Contexts and Academic Competitions	Competitions for Competency Alignment	Alignment with the "Lanqiao Cup" competition competencies; integration of

		competition and classroom instruction
Single-dimensional evaluation methods that overlook the learning process	Holistic Assessment System	Five-Dimensional Formative Assessment

2.3. Progressive Project Chain

This course restructures 16 academic hours of hands-on instruction into four progressively advancing projects, each lasting 4 academic hours. The four projects unfold step-by-step along the path of "Modeling from Scratch—Transfer Learning—Few-Shot Fine-Tuning—Competition Practice." Each project begins with a driving problem closely aligned with real-world engineering scenarios and progresses incrementally across three dimensions: knowledge, data scale, and competition capabilities, as shown in Table 2.

Project 1: Modeling from Scratch. Driven by embedded, low-computing-power scenarios, this project requires students to design a lightweight CNN from scratch on edge devices where pre-trained models cannot be used. It focuses on developing an awareness of parameter control and involves completing data loading, training loops, and confusion matrix analysis to establish a complete cognitive loop encompassing "data—model—training—evaluation."

Project 2: Pre-trained Model Transfer. Driven by the scenario of "a company needing to rapidly deploy a high-precision classification service," students train a ResNet-18 classification head while freezing the convolutional layers. They then conduct a quantitative comparison with Project 1 using accuracy curves and t-SNE feature visualizations. Compared to Project 1, this project represents a progression from building models from scratch to comparing different approaches, fostering an awareness of quantitative comparison based on experimental data.

Project 3: Few-Shot Fine-Tuning. Focusing on the common challenge of few-shot learning in real-world industries, students trained high-precision classifiers using techniques such as data augmentation, differential learning rates, and early stopping mechanisms, and conducted multidimensional comparisons with models trained from scratch. Compared to Project 2, this project advanced by introducing more refined tuning methods used in engineering practice, directly addressing the few-shot classification tasks commonly found in the Lanqiao Cup.

Project 4: Competition Practice. Using a multi-class, fine-grained recognition task modeled after Kaggle competitions, students comprehensively apply what they have learned in the first three projects to generate predictions through ensemble voting of models with multiple random seeds. They are also encouraged to implement Grad-CAM visualization and simple web deployment. Compared to Project 3, this project represents a progression from individual techniques to comprehensive practical application, emphasizing both submittable results and explainable processes.

Table 2. Progressive Comparison of the Four Projects

Dimension	Project 1	Project 2	Project 3	Project 4
Project Theme	Modeling from Scratch	Pre-training Transfer	Fine-Tuning with Few Samples	Competition Practice
Application Scenarios	Embedded Systems with Limited Computing Power	Rapid Service Deployment	Quick-and-Dirty Small-Sample Tasks	Kaggle Competitions
Progressive Learning	Custom CNN	Pre-trained ResNet-18	Differential Fine-Tuning	Model Ensemble
Data Scale	CIFAR-10	CIFAR-10 + ImageNet pre-training	Hot Dog Dataset	Fine-grained dog breeds
Key Technologies	Lightweight CNNs, Parameter Control	Freezing Convolutional Layers, t-SNE	Data Augmentation, Early Stopping	Ensemble Voting, Grad-CAM
Competition Performance Benchmarking	Foundational Modeling	Solution Comparison	Handling Small Datasets	Comprehensive Practical Application

2.4. Digital and Intelligent Tools

This course builds a smart teaching environment that is AI-assisted, human-led, and integrates virtual and real-world elements. During the code development phase, AI programming assistants such as GitHub Copilot and Cursor are introduced. Students are required to follow a standardized workflow: requirements specification → AI-generated code → human review → iterative optimization. Directly copying and pasting AI-generated code is strictly prohibited, with a focus on cultivating the ability to precisely articulate requirements and review code. In the learning support phase, students are encouraged to use large language models for concept explanation and error diagnosis; however, they must document the complete question–answer–verification chain. By comparing outputs from multiple models, students develop information discernment skills and learn to proactively identify and correct "hallucinations" in large language models. In the model understanding phase, tools such as TensorBoard, Netron, and Grad-CAM are used to enable real-time monitoring of training, visualization of network structures, and presentation of decision-making rationale, helping students overcome difficulties in understanding the "black-box" nature of deep learning.

2.5. Formative Assessment Throughout the Course

To address the limitations of traditional, one-dimensional assessment methods, this course has established a system of process-based formative assessment and multidimensional comprehensive evaluation. The final grade is composed of five components: project process performance (30%), proficiency in using AI tools (15%), technical competence (25%), summative defense (20%), and personal growth portfolio (10%). Phased assessments for each project (coding standards,

model accuracy, report quality, etc.) are evaluated using project-level rubrics and serve as the primary basis for corresponding course-level assessments (see Table 3). This integrates process-based performance and technical outcomes into a single framework, facilitating a shift from evaluating isolated outcomes to assessing developmental competencies.

Table 3. Course-Level Evaluation Dimensions and Data Sources

Course-Level Dimensions (Weight)	Primary Data Sources
Project Process Performance (30%)	Project milestone submission records, Git commit logs
AI Tool Literacy (15%)	Reflection logs on AI usage
Technical Competence (25%)	Technical Scores for Each Project (Code Quality, Architecture, Rigor of Experiments)
Final Presentation (20%)	Project 4 Team Defense
Personal Growth Portfolio (10%)	Self-assessments of competencies at the beginning, middle, and end of the semester

3. Teaching Implementation and Outcomes

3.1. Teaching Implementation

This course was implemented during the first semester of the 2026–2027 academic year in a pilot class for the Intelligent Science and Technology major at a certain university, with a total of 42 students divided into groups of 4–5. The course utilized a cloud-based GPU training environment, where students completed development using Jupyter Notebooks and submitted their progress through Git. The four projects, in sequence, were modeling from scratch, transfer learning, few-shot fine-tuning, and competition practice. To track teaching effectiveness, self-assessment questionnaires were distributed before and after the course, and upon project completion, compliance with standards was evaluated through a combination of automated assessment and manual review.

During the course, instructors intervened to address two common issues. First, some students submitted AI-generated code directly at the beginning of a project without first understanding or reviewing it. Instructors required them to redo the work following the process of "requirement description→AI generation→ manual review→iterative optimization", and this process was recorded as part of their AI tool literacy assessment. Second, some students became bogged down in blind trial-and-error during parameter tuning for Project 2. Instructors guided them back to an experimental approach involving fixed variables and item-by-item comparison—first clarifying the objects of comparison, then adjusting a single variable. These interventions prompted students to shift from merely completing experiments to purposeful training in engineering decision-making.

3.2. Data Sources

This study collected three types of data based on the teaching implementation process in the pilot class. The first consists of project completion data for the four projects, obtained through a combination of automated assessment and manual review, which reflects the status of project completion. Second, student self-assessment questionnaires administered before and after the course were designed around five dimensions—including independently completing project workflows, making technical selection judgments, and the proper use of AI tools—comprising a total of 20 items. Using a 5-point Likert scale, 42 questionnaires were distributed, 42 were returned, and 42 were valid, resulting in a 100% response rate and a 100% validity rate. The overall Cronbach's α coefficient for the questionnaire was 0.913, indicating good reliability. Third, the registration and preliminary round results for the Lanqiao Cup AI competition were used to observe the synergy between the classroom project and the competition.

Given that the course is still in the early stages of reform and the sample size is limited, the effectiveness analysis primarily relied on descriptive statistics; paired-sample t-tests were used for pre- and post-test self-assessment data, and the Shapiro-Wilk test was employed to assess the normality of the pre- and post-test differences.

3.3. Project Achievement Status

After the course concluded, the compliance status of the four projects was summarized based on the results of automated evaluation and manual review (Table 4). All four projects met the predetermined passing requirements, and the compliance rate showed an overall upward trend as the projects progressed. From Project 1 to Project 4, the passing standard gradually increased from a test accuracy rate of no less than 60% to an ensemble accuracy that was at least 3 percentage points higher than that of the best single model; accordingly, the compliance rate rose from 73.8% to 92.9%. In particular, after introducing model assembling in Project 4, the ensemble accuracy of most groups improved by an average of 3–5 percentage points compared to the best single model, and the overall task completion rate was higher than that of the first three individual projects. Since the passing criteria and task difficulty varied across projects, the longitudinal changes in the compliance rates merely reflect the completion status of each project and should not be directly interpreted as evidence of a step-by-step improvement in student ability.

Table 4. Comparison of Pass Rates Across the Four Projects

Project	Passing Standard	Number of Students Meeting the Standard	Pass Rate	Sample Size N
Project 1	Test Accuracy ≥ 60%	31	73.8%	42
Project 2	Test accuracy ≥ 70%	34	81.0%	42
Project 3	Test accuracy ≥ 85%	36	85.7%	42
Project 4	Ensemble accuracy	39	92.9%	42

	improved by ≥3 percentage points			

3.4. Student Competency Development

A comparison of student self-assessment questionnaires before and after the course implementation revealed significant improvements in 3 out of 5 dimensions, while the other 2 dimensions showed no significant changes. In the dimension "Independently completing the entire deep learning project process," self-assessment scores increased from 2.86 ± 0.91 on the pre-test to 4.12 ± 0.73 on the post-test; in the dimension "Ability to make technical selection decisions through comparative experiments," the score increased from 2.74 ± 0.88 to 3.95 ± 0.81; and in the dimension "Ability to conduct standardized development using AI tools," the score increased from 2.52 ± 0.96 to 4.21 ± 0.69. According to the Shapiro-Wilk test, the pre- and post-test differences across all three dimensions conformed to a normal distribution; therefore, a paired t-test was conducted. The results showed that the differences between the pre- and post-tests for all three dimensions reached statistical significance ($t = 7.12, 6.85, 8.03$; $P < 0.001$ for all). These results preliminarily demonstrate the positive role of progressive project design in cultivating students' systems engineering thinking, their ability to make informed technology selection decisions, and their proficiency in the proper use of AI tools.

3.5. Competition Integration

The pilot class for this course consisted of 42 students, 32 of whom registered for the Lanqiao Cup Artificial Intelligence competition. Twenty-one students passed the preliminary round, resulting in a pass rate of 65.6%, which is higher than the rate for the same period in previous years when this project-based reform had not yet been implemented (approximately 38%). Since the two groups serve as non-contemporaneous historical controls and their comparability has not been verified, this difference is for reference only. From the perspective of skill transfer, there is a correspondence between the competition tasks and the classroom projects in terms of problem structure: Projects 1 and 2, which train students in data loading and model building, correspond to the basic modeling phase of the competition; Project 3, which focuses on data augmentation and differential fine-tuning, corresponds to the common small-sample processing tasks in the competition; and Project 4, which covers model ensemble and result submission, directly corresponds to the competition's submission format requirements. Some students reported that the thinking patterns developed in the classroom projects—such as breaking down driving problems, designing solutions, and conducting comparative validation—could be directly transferred to the process of solving competition problems, providing preliminary evidence supporting the design intent of integrating competitions and classroom instruction.

3.6. Limitations of the Study

All findings in this study are derived from a single pilot class; the sample size, control design, and methods of effect measurement are all relatively limited. Pass rates are influenced by varying passing standards across different projects, making longitudinal comparisons difficult; competency development relies on student self-assessment, which is subject to subjective bias; and competition pass rates are based on asynchronous historical references, resulting in insufficient comparability. Therefore, the above results should be regarded as exploratory evidence that preliminarily supports the feasibility of an incremental project design and the integration of project-based learning with teaching competitions; however, causal inferences cannot yet be drawn from them.

4. Conclusion

Addressing issues such as fragmented deep learning practice courses, a disconnect between classroom instruction and competitions, and one-dimensional assessment methods, this paper restructured a four-stage progressive project chain, established a digital and intelligent tool support system and a comprehensive assessment framework, and conducted exploratory practices in a single pilot class. Preliminary results show that all four-stage projects met the predetermined passing requirements; students' self-assessment scores significantly improved in project workflow, technology selection, and the proper use of AI tools; and the aligned design of classroom projects and competition skills was preliminarily validated. These results tentatively support the feasibility of integrating progressive project design with academic competitions; however, their generalizability remains to be verified through larger-scale, multi-cycle-controlled studies.

Acknowledgements

This work was supported by the 2025 Shanghai Higher Education Young Teachers Training and Funding Program, and the 2026 Project-based Curriculum Construction Project of Shanghai Lixin University of Accounting and Finance.

References

[1] Xu, M. Lu, S.W. and Yang, B. (2026) Teaching Practices in Digital Image Processing Driven by Deep Learning. Journal of Hubei Normal University (Natural Science Edition), 46 (3): 104-110.

[2] Zhai, M.L. (2026) A Study on Innovative Teaching Methods for the Data Structures Course Under the "Human-Machine Dual-Instructor Collaboration" Model. Internet of Things Technology, 16 (20): 145–149.

[3] Deng, W.S. (2026) A Study on Enhancing Students' Deep Learning Abilities in the Context of Integrating Ideological and Political Education into Blended Learning. University, S1: 53–55.

[4] Cheng, Y. and Li, F. (2024) Exploration and Practice of a Project-Driven Teaching Model for the "Introduction to Deep Learning" Course, Higher Education in Chemical

Engineering, 41(6): 37-44.

[5] Li, X. Lu, X. and Liu, F. (2025) Reform of Computer Vision Courses Based on Project-Driven and Interdisciplinary Integration: A Case Study of Semantic Segmentation of Remote Sensing Images, Software Guide, 24(11): 67-75.

[6] Gao, Y. Lin, Y. and Wu, D.W. et al. (2021) Exploration of Project-Driven, Blended Online-Offline Heuristic Deep Learning Practical Instruction. Computer Education, 9: 51–53, 57.

[7] Ma, Y.K. Li, X.W. and Gu, L.S. et al. (2026) Exploring a Research-Based Teaching Model for Deep Learning Courses, Journal of Higher Education, 6: 129-132.

[8] Hu, X. Pei, L.L. Yuan, B. et al. (2026) Practice of an Artificial Intelligence Teaching System Aimed at Cultivating Practical Innovation Abilities, Computer Education, 6: 192–197.

[9] Li, W.F. Fan, H.J. Tian, H.J. et al. (2024) "Tiered, Progressive Programming Laboratory Instruction Integrating Teaching, Practice, and Competition", Computer Education, 1: 204–208.