УДК 004.896:004.93:621.3.049.75 **DOI:** 10.31673/2412-9070.2026.056302

А. А. ВЕРЛАНЬ[12], д-р техн. наук, професор;
ORCID: 0000-0002-6469-2638
Є. Р. ФАСТОВЕЦЬ[1], аспірант,
ORCID: 0009-0003-5317-9518
[1]Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
[2]Норвезький університет природничих та технічних наук

БАЛАНСУВАННЯ «ДОВГОГО ХВОСТА» ДЛЯ ПОКРАЩЕННЯ ПОВНОТИ ВИЯВЛЕННЯ РІДКІСНИХ КЛАСІВ БЕЗ ДЕГРАДАЦІЇ ТОЧНОСТІ ПРИ ДЕТЕКЦІЇ ДЕФЕКТІВ ПЕЧАТНИХ ПЛАТ

У статті розглянуто задачу балансування «довгого хвоста» у системах автоматизованого оптичного контролю друкованих плат, Automated Optical Inspection (AOI) для Printed Circuit Board (PCB). Такий розподіл є типовим для виробничих даних: кілька видів дефектів трапляються часто, тоді як рідкісні, але технологічно важливі класи мають малу кількість прикладів у навчальній вибірці. Унаслідок цього нейромережевий детектор краще навчається на частих класах і гірше виявляє дефекти хвоста розподілу. Для контролю якості друкованих плат ця проблема є суттєвою, оскільки пропуск реального дефекту може призвести до відмови виробу, а неконтрольоване збільшення кількості хибнопозитивних спрацьовувань підвищує навантаження на оператора та зменшує пропускну здатність виробничої лінії.

Основна ідея дослідження полягає не в простому збільшенні повноти виявлення, тобто recall, а в її покращенні для рідкісних класів за умови збереження прийнятної точності позитивних спрацьовувань, precision і контрольованої кількості False Positive (FP) на одну плату. У роботі запропоновано методичну схему застосування навчальних і постобробних підходів: контрольованого семплінгу, hard negatives, Focal Loss, Class-Balanced Loss, Equalization Loss v2, Seesaw Loss, двостадійного навчання, калібрації та підбору класових порогів. Окрему увагу приділено протоколу оцінювання, що включає recall за заданого рівня precision, FP/плата, macro-recall, per-class precision-recall криві та перевірку затримки інференсу.

Практична цінність роботи полягає у формуванні керованої схеми налаштування AOI-детектора, у якій покращення якості на рідкісних дефектах не розглядається ізольовано від виробничих обмежень і ризику деградації precision.

Ключові слова: AOI, PCB, друковані плати, довгий хвіст, дисбаланс класів, нейромережева детекція, recall, precision.

Вступ

Автоматизований оптичний контроль друкованих плат є одним із важливих етапів контролю якості у виробництві електроніки. AOI-система перевіряє плату перед наступними технологічними операціями або перед відвантаженням, тому помилки детектора мають не тільки технічні, а й економічні наслідки. Пропуск реального дефекту, False Negative (FN) може призвести до відмови виробу, рекламацій або додаткових витрат на пізніх етапах виробництва. Хибне спрацювання, FP, також небажане: воно збільшує кількість ручних перевірок, навантажує оператора і знижує пропускну здатність лінії.

На практиці AOI-система має працювати не просто з високим середнім значенням mean Average Precision (mAP). Для виробництва важливо, щоб критичні дефекти не пропускалися, а кількість хибних спрацьовувань залишалася у допустимих межах. Тому, для задачі контролю

© Верлань А. А., Фастовець Є. Р., 2026

якості доцільно розглядати не лише загальні метрики, а й recall рідкісних класів за умови заданого precision та обмеження на середню кількість хибнопозитивних спрацьовувань на одну плату, далі FP/плата.

Постановка задачі

Окрема складність виникає через довгохвостий розподіл дефектів. Частина класів трапляється часто і формує «голову» розподілу. Інші класи зустрічаються рідко, хоча саме вони можуть бути важливими з погляду надійності. Якщо нейромережевий детектор, зокрема модель на основі згорткових нейронних мереж, Convolutional Neural Networks (CNN), навчається на такій вибірці без додаткових механізмів балансування, то оптимізація загальної функції втрат переважно підлаштовується під часті класи. У результаті head-класи мають прийнятну якість, а tail-класи залишаються слабкими.

Очевидне рішення, наприклад, багаторазове повторення рідкісних прикладів у навчанні, не завжди працює стабільно. Сильний oversampling може призвести до перенавчання на невеликій кількості хвостових зразків. Модель починає бачити дефект там, де є лише відблиск, текстура міді, край доріжки або інший схожий фоновий патерн. Через це зростає FP, а precision падає. Отже, задача полягає не в тому, щоб будь-якою ціною покращити recall хвоста, а в тому, щоб зробити це під явним контролем precision і FP/плата.

Аналіз останніх досліджень і публікацій

У задачах щільної детекції об’єктів Focal Loss зменшує внесок легких прикладів і концентрує навчання на складних випадках. Для AOI це важливо, бо фон часто займає більшу частину зображення, а сам дефект може бути малим і локальним [1]. Class-Balanced Loss вводить ваги на основі ефективної кількості зразків у класі. Такий підхід м’якший за пряме зважування за правилом 1/f і менше ризикує зруйнувати якість на частих класах [2].

Для довгохвостої класифікації показано, що представлення ознак і класифікатор не обов’язково потрібно навчати в однаковому режимі. Підхід із декуплінгом representation/classifier дозволяє спочатку навчити стабільні ознаки на природному розподілі, а потім окремо коригувати межі рішень для класів [3]. Для AOI це корисно, бо надто агресивне балансування на ранніх етапах може змусити модель реагувати на фонові структури плати як на дефекти.

У long-tailed object detection проблема часто пов’язана не лише з кількістю прикладів, а й з дисбалансом градієнтів. Рідкісні класи отримують мало позитивних оновлень, тоді як негативні градієнти від частих класів і фону можуть переважати. Equalization Loss v2 (EQLv2) пропонує керувати балансом позитивних і негативних градієнтів для різних категорій [4]. Seesaw Loss використовує механізми mitigation і compensation, щоб послаблювати небажаний тиск head-класів на tail-класи та водночас стримувати хибнопозитивні спрацьовування [5].

Окремий напрям стосується семплінгу. Instance-Aware Repeat Factor Sampling (IRFS) враховує кількість інстансів, а не лише кількість зображень, і тому краще підходить для детекції, де одне зображення може містити різну кількість дефектів [6]. Також використовуються методи корекції логітів. Logit adjustment змінює вихідні оцінки моделі з урахуванням апріорних частот класів [7]. Balanced Meta-Softmax або Balanced Softmax враховує зсув між навчальним і тестовим розподілами міток [8]. FRACAL пропонує посткалібрацію для long-tailed detection, використовуючи фрактальну розмірність як ознаку просторового розподілу класів [9].

У сфері PCB існують відкриті набори даних, зокрема DeepPCB [10] і DsPCBSD+ [11]. Вони дають базу для порівняння методів, але реальні виробничі дані часто мають сильніший перекіс між класами. Крім того, в AOI присутні доменні фактори: відблиски, нерівномірне освітлення, різна текстура міді, зміни камери та партії виробництва. Саме ці фактори часто провокують FP, тому методи балансування треба оцінювати не тільки за приростом recall, а й за стабільністю precision.

Мета і задачі дослідження

Метою роботи є обґрунтування методичного підходу до покращення повноти виявлення, тобто recall, для рідкісних класів дефектів PCB без деградації precision і без неконтрольованого зростання кількості хибнопозитивних спрацьовувань на плату.

Для досягнення цієї мети розглянуто практичну схему навчання і налаштування one-stage детекторів. До неї входять контрольований семплінг, функції втрат для дисбалансованих даних, двостадійне навчання, калібрація, підбір порогів для окремих класів і протокол перевірки, який враховує як якість виявлення рідкісних дефектів, так і виробниче навантаження від FP.

Виклад основного матеріалу

Основна частина статті побудована навколо практичної схеми роботи з довгохвостим розподілом дефектів у задачах AOI PCB. Спочатку розглядаються метрики та правила розбиття даних, далі подано групи методів балансування рідкісних класів, підходи до калібрації порогів, протокол оцінювання та рекомендації щодо впровадження в AOI-систему.

Метрики та розбиття даних

Для задачі AOI недостатньо контролювати тільки загальний mAP. Якщо модель добре знаходить часті дефекти, але пропускає рідкісні критичні класи, середня метрика може виглядати прийнятно, хоча реальна користь такої моделі буде обмеженою. Тому для tail-класів основними показниками доцільно вважати Recall@Precision, FP/плата, macro-recall і per-class precision-recall (PR) криві.

Recall за заданого рівня precision показує, наскільки можна підвищити повноту виявлення без падіння точності позитивних спрацьовувань нижче прийнятного рівня. FP/плата пов'язує якість детектора з виробничим навантаженням на оператора. Macro-recall не дає частим класам повністю домінувати в оцінці, тому краще відображає поведінку на хвості.

Оцінювання має спиратися на коректне розбиття даних. Для PCB небезпечним є витік майже однакових плат між train і validation. Якщо схожі зображення однієї плати або однієї партії потрапляють у різні частини вибірки, якість буде завищеною. Тому спліти бажано формувати за партіями виробництва, типами плат, сесіями зйомки або іншими ознаками, які відокремлюють близькі зразки.

Для рідкісних класів корисно мати окремий stress-test набір. До нього варто включати приклади з низьким контрастом, відблисками, малими дефектами, складним фоном і неоднозначними межами дефекту. Такий набір не замінює основну валідацію, але допомагає побачити, чи не тримається покращення тільки на простих прикладах.

Методи балансування рідкісних класів
Семплінг, синтетичні приклади та hard negatives

Щоб рідкісні класи частіше потрапляли в батчі, можна використовувати Repeat Factor Sampling або IRFS [6]. У PCB-детекції це треба робити обережно. Якщо просто багато разів дублювати хвостові приклади, модель швидко запам'ятає конкретні фрагменти плат і почне переносити ці ознаки на схожі, але нормальні області. Тому для repeat sampling бажано вводити верхню межу повторення і перевіряти вплив кожного рівня oversampling на FP/плата.

Hard negatives є важливою частиною такого навчання. Під hard negatives маються на увазі зразки або області, на яких модель помилково бачить дефект. Для AOI це можуть бути відблиски, технологічні отвори, краї доріжок, залишки флюсу, текстура міді або інші елементи, які візуально схожі на дефект. Додавання таких прикладів у навчання допомагає піднімати recall хвоста без неконтрольованого зростання FP.

Синтетичні приклади також можуть бути корисними, але вони не повинні створювати помітний доменний зсув. Найбезпечніші варіанти для PCB: copy-paste дефектних патчів у реалістичні області плати, легка зміна яскравості та контрасту, локальний шум, малі афінні деформації. Відблиски або смуги краще моделювати не лише як позитивні приклади, а й як негативні, якщо в реальному виробництві вони часто провокують FP.

Частку синтетики доцільно контролювати окремо. Якщо синтетичні зразки не відмічати, під час аналізу помилок буде складно зрозуміти, чи модель навчилася на справжніх дефектах, чи на штучних артефактах генерації. Тому для таких прикладів бажано зберігати окремий тег або хоча б вести журнал їхнього походження.

Функції втрат і баланс градієнтів

Focal Loss [1] добре підходить для ситуацій, коли фон домінує над об'єктами, а легких негативних прикладів набагато більше, ніж складних. У PCB це типова ситуація: більша частина зображення є нормальною областю плати, а дефект займає невеликий фрагмент. Параметр γ у Focal Loss визначає, наскільки сильно зменшується внесок легких прикладів. Його не варто підбирати лише за mAP, бо надто агресивне фокусування може покращити recall хвоста і водночас збільшити FP.

Class-Balanced Loss [2] використовує ваги, пов'язані з ефективною кількістю зразків у класі. На відміну від прямого зважування за частотою, цей підхід не так різко підсилює дуже малі класи. Для PCB це важливо, бо деякі рідкісні дефекти можуть мати лише кілька прикладів, і надмірна вага для них зробить навчання нестабільним. Параметр β слід підбирати за метриками Recall@Precision і FP/плата.

У long-tailed detection проблема часто проявляється як нерівний потік градієнтів між класами. Часті класи багато разів беруть участь у навчанні, а хвостові класи отримують мало позитивних оновлень. EQLv2 [4] намагається вирівняти цей процес через керування негативними градієнтами. Seesaw Loss [5] можна використати як альтернативу, особливо якщо після ребалансування помітно зростає кількість хибнопозитивних спрацьовувань.

Для AOI важливо перевіряти не тільки приріст recall, а й те, за рахунок чого він отриманий. Якщо модель починає частіше реагувати на фон, формальне покращення хвостових класів буде нестійким. Тому після кожної зміни функції втрат треба дивитися per-class PR-криві, FP/плата і приклади помилок.

Декуплінговане навчання

Ідею розділення навчання представлення і класифікатора [3] можна перенести в AOI як двостадійний процес. На першому етапі, Stage A, повний детектор навчається на доменно реалістичних даних із помірними аугментаціями. На цьому етапі не слід занадто сильно ребалансувати хвіст, бо backbone має навчитися стабільних ознак плати, дефектів і фону.

На другому етапі, Stage B, backbone фіксується або донавчається з дуже малим learning rate (LR), а head адаптується на ребалансованому семплінгу чи зважених функціях втрат. Такий підхід зменшує ризик, що модель почне сприймати нормальні структури PCB як дефекти. Фактично Stage B коригує межі рішень, не руйнуючи вже вивчене представлення.

Якщо FP є особливо дорогими, задачу можна розділити на два рівні. Спочатку детектор формує кандидати «дефект/фон», а потім окремий класифікатор визначає тип дефекту на кропі. У такій схемі довгохвосте ребалансування переноситься переважно на класифікатор кропів, де менше фонових негативів. Для цього рівня простіше застосовувати Balanced Softmax [8] або logit adjustment [7].

Калібрація порогів і контроль FP/плата

Навіть добре навчений детектор може мати некалібровані оцінки впевненості. Через це два класи з однаковим confidence score можуть мати різну фактичну ймовірність правильного спрацьовування. У задачі з довгим хвостом це особливо помітно: для рідкісного класу модель може бути менш упевненою навіть тоді, коли кандидат справді відповідає дефекту.

Тому налаштування порогів для окремих класів доцільно розглядати як окремий етап постобробки, а не як механічне зниження єдиного порога для всієї моделі. Для кожного tail-класу можна підібрати поріг, який покращує recall і водночас не опускає precision нижче прийнятного рівня. На валідації для цього будують PR-криві, після чого окремо перевіряють сумарну кількість FP на плату. Це потрібно тому, що одночасне зниження порогів для кількох рідкісних класів може виглядати прийнятним для кожного класу окремо, але створити надмірне навантаження на перевірку плати в цілому.

Пост-хок калібрація корисна тоді, коли модель уже навчена або розгорнута, а повне перевчання є дорогим. У такому випадку можна застосовувати logit adjustment [7], температурну калібрацію або FRACAL [9]. У YOLO-подібних пайплайнах стандартна конфігурація зазвичай задає глобальний confidence threshold і поріг Intersection over Union (IoU) для Non-Maximum

Suppression (NMS) [12]. Якщо потрібні per-class thresholds, їх доцільно реалізовувати окремою постобробкою або модифікацією інференсу, оскільки один глобальний поріг не завжди достатній для виробничої AOI-задачі.

Протокол оцінювання та активне донавчання

Щоб результат не залежав від випадково вдалого спліту, протокол експериментів треба фіксувати заздалегідь. До нього входять політика аугментацій, правила семплінгу, параметри NMS, вхідна роздільна здатність, допустимий рівень FP/плата і спосіб підбору класових порогів. Без цього складно зрозуміти, яка саме зміна дала приріст і чи не був він отриманий ціною деградації precision.

Абляції доцільно проводити послідовно. Спочатку оцінюється базова модель. Потім додається IRFS або інший контрольований семплінг. Далі перевіряються Focal Loss і Class-Balanced Loss, після цього EQLv2 або Seesaw Loss, Stage B, класові пороги та калібрація. Кожен крок слід оцінювати за однаковими метриками: recall за заданого рівня precision, FP/плата, macro-recall, per-class PR і latency.

Після впровадження модель не повинна залишатися статичною. У виробництві змінюються партії плат, освітлення, камера, матеріали й технологічні допуски. Через це треба збирати приклади для активного донавчання. Найціннішими є кадри з низькою впевненістю, повторюваними FP на однакових структурах і випадки, де оператор не погодився з рішенням моделі. Такі приклади можна періодично додавати до hard negatives або використовувати для короткого Stage B донавчання.

Паралельно слід моніторити частоту FP, розподіл confidence scores і частоту детекцій по класах. Різка зміна цих показників часто вказує не на покращення або погіршення самої моделі, а на дрейф домену: зміну освітлення, положення камери, якості поверхні або типу плати.

Порівняння груп методів

У табл. 1 узагальнено групи методів, які доцільно перевіряти для підвищення tail-recall у задачах AOI PCB. У таблиці спеціально зазначено не тільки очікуваний позитивний ефект, а й основний ризик для precision, оскільки саме цей компроміс є центральним для запропонованої постановки задачі.

Таблиця 1

Порівняння стратегій покращення tail-recall у задачах AOI PCB

Група методів	Очікуваний вплив на tail-recall	Основний ризик і контроль
Семплінг, IRFS/RFS [6]	Рідкісні інстанси частіше потрапляють у навчання.	Перенавчання на хвості; потрібні cap для повторів і hard negatives.
Focal Loss, Class-Balanced Loss [1], [2]	Фокус на складних прикладах і м'якше міжкласове балансування.	Зростання FP; потрібен підбір γ, β і контроль фону.
EQLv2, Seesaw Loss [4], [5]	Балансування градієнтів для rare-класів.	Можлива нестабільність між класами; потрібні абляції.
Декуплінг, Stage A/B [3]	Підняття хвоста без руйнування ознак backbone.	Недонавчання head; потрібний обережний fine-tuning.
Калібрація і класові пороги, FRACAL [9], [12]	Приріст recall без повного перевчання.	Накопичення FP; контроль через precision і FP/плата.

У табл. 2 наведено метрики, за якими варто контролювати якість роботи детектора під час валідації. Такий набір показників дає змогу відокремити реальне покращення хвоста від ситуації, коли приріст recall отримано ціною надмірної кількості хибних кандидатів.

Таблиця 2

Метрики для контролю обмежень у валідації

Метрика	Що показує	Як використовується
Recall@Precision	Recall за прийнятного рівня precision.	Основний KPI для рідкісних класів.
FP/плата	Середня кількість хибних кандидатів на плату.	Контроль навантаження на оператора.
Macro-recall	Середнє recall по класах.	Чутливий показник для хвоста.
Per-class PR	Precision-recall криві для кожного класу.	Підбір класових порогів і пошук слабких класів.
Latency	Час обробки однієї плати.	Перевірка виробничих обмежень.

Практичні налаштування one-stage детекторів

Для дрібних дефектів PCB часто корисно збільшувати вхідну роздільну здатність або використовувати tile/patch-стратегію. Це допомагає зберегти локальні деталі, які можуть зникати при сильному масштабуванні зображення. Водночас збільшення роздільної здатності підвищує latency, тому його треба оцінювати разом із виробничими обмеженнями. У практичних AOI-системах це особливо важливо для YOLO-подібних моделей, You Only Look Once (YOLO), які часто обирають через швидкість інференсу.

Аугментації потрібно підбирати з урахуванням фізичного змісту зображення. Наприклад, brightness/contrast, Gaussian noise і малі афінні перетворення зазвичай не руйнують структуру плати. Mosaic може бути корисним у загальних задачах object detection, але для AOI він інколи створює неприродний контекст і змішує фрагменти плат так, як це не трапляється на виробництві. Тому його частку бажано зменшувати або перевіряти окремою абляцією.

Окремої уваги потребує NMS. Якщо поріг для хвостового класу знижено, модель може формувати кілька близьких боксів із невисокою впевненістю. У такому випадку треба перевіряти IoU-поріг NMS і, за потреби, використовувати class-aware NMS для критичних класів. Якщо на різних виробничих лініях використовуються різні камери або режими освітлення, доцільно мати окремі параметри калібрації та класові пороги для кожного домену. Це простіше, ніж підтримувати кілька незалежних моделей, і водночас зменшує ризик падіння precision.

Пояснюваність, розмітка хвоста і впровадження

Коли після ребалансування зростає кількість FP, корисно застосовувати методи пояснюваності. Для класифікатора кропів можна використовувати Gradient-weighted Class Activation Mapping (Grad-CAM). Такий аналіз показує, на які області зображення модель спирається під час прийняття рішення. Якщо активація зосереджена не на дефекті, а на відблиску, краї плати або текстурі міді, це сигнал для доопрацювання даних.

Розмітку хвостових класів бажано організовувати не випадково, а за категоріями ризику. Спочатку варто збирати дефекти з найбільшою вартістю FN. Далі можна додавати класи, які модель часто плутає між собою, і типові FP-патерни. Такий підхід економить час розмітки, бо оператор працює не з усім потоком даних, а з прикладами, які найбільше впливають на якість моделі.

Узагальнений порядок впровадження може бути таким. Спочатку визначають множину класів, tail-підмножину, прийнятний рівень precision, допустимий рівень FP/плата і відносну критичність помилок. Потім формують коректні train/validation/test спліти, навчають базову модель, додають контрольований семплінг і hard negatives, перевіряють Focal/Class-Balanced

втрати, підключають EQLv2 або Seesaw Loss, виконують Stage B, підбирають класові пороги, проводять калібрацію і запускають моніторинг. Така послідовність робить систему керованою: кожен крок можна перевірити окремо і відкотити, якщо він погіршує precision або FP/плата.

Обговорення результатів

Запропонований підхід не замінює емпіричне тестування, але задає практичну рамку, у якій можна оцінювати методи для довгохвостого AOI. Його перевага в тому, що він одразу враховує виробничі обмеження. Для практичної системи важливо не просто збільшити кількість знайдених рідкісних дефектів, а зробити це так, щоб оператор не отримав надмірний потік хибних спрацьовувань.

Основний компроміс виникає між recall хвоста і precision. Семплінг, ваги втрат і зниження порогів можуть піднімати recall, але кожен із цих методів здатний збільшити FP. Саме тому їх не слід оцінювати ізольовано. Якщо метод дає приріст macro-recall, але одночасно виводить FP/плата за допустиму межу, він не підходить для виробничої AOI-системи без додаткової калібрації або збору hard negatives.

Ще одна важлива умова полягає в коректності валідації. Якщо validation набір не відображає реальні зміни освітлення, партій і типів плат, то підібрані класові пороги можуть бути нестабільними. Тому запропонований підхід має використовуватися разом із доменно правильним протоколом оцінювання.

Висновки та перспективи подальших досліджень

У статті запропоновано методичний підхід до балансування «довгого хвоста» в AOI PCB. Основний акцент зроблено на покращенні recall рідкісних класів без деградації precision і без неконтрольованого збільшення кількості FP на одну плату. Така інтерпретація краще відповідає виробничій логіці AOI, де важливі не лише середні метрики детектора, а й реальне навантаження на процес контролю якості.

На основі цього підходу обґрунтовано поєднання кількох груп методів: контрольованого семплінгу, hard negatives, Focal/Class-Balanced втрат, EQLv2 або Seesaw Loss, двостадійного навчання Stage A/B, класових порогів і пост-хок калібрації. Основна ідея полягає не в загальному підвищенні mAP, а в керованому покращенні recall хвостових класів із постійним контролем precision і FP/плата.

Подальші дослідження доцільно спрямувати на генерацію реалістичніших синтетичних дефектів, кращу інтеграцію Grad-CAM у цикл аналізу FP, а також на виявлення доменного дрейфу під час тривалої експлуатації AOI-системи.

Внесок авторів

Андрій ВЕРЛАНЬ – наукове керівництво, уточнення постановки задачі; Євгеній ФАСТОВЕЦЬ – постановка задачі, аналіз останніх досліджень і публікацій, виклад основного матеріалу.

Декларація про штучний інтелект

Під час підготовки рукопису інструменти на основі штучного інтелекту могли використовуватися для узгодження термінології. Наукова ідея, постановка задачі, добір джерел, інтерпретація положень і відповідальність за зміст статті належать автору.

Конфлікт інтересів

Автор заявляє про відсутність конфлікту інтересів та підтверджує, що під час підготовки цієї роботи не існувало комерційних, фінансових чи інших взаємовідносин, які могли б вплинути на результати дослідження або їх інтерпретацію. Робота виконана відповідно до принципів академічної доброчесності, етичних норм проведення наукових досліджень та вимог редакційної політики щодо запобігання конфлікту інтересів.

Список використаної літератури

1. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 2980-2988). Retrieved from https://doi.org/10.1109/ICCV.2017.324

2. Cui, Y., Jia, M., Lin, T.-Y., Song, Y., & Belongie, S. (2019). Class-balanced loss based on effective number of samples. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 9268-9277). Retrieved from https://doi.org/10.1109/CVPR.2019.00949

3. Kang, B., Xie, S., Rohrbach, M., Yan, Z., Gordo, A., Feng, J., & Kalantidis, Y. (2020). Decoupling representation and classifier for long-tailed recognition. International Conference on Learning Representations. Retrieved from https://doi.org/10.48550/arXiv.1910.09217

4. Tan, J., Lu, X., Zhang, G., Yin, C., & Li, Q. (2021). Equalization Loss v2: A new gradient balance approach for long-tailed object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1685-1694). Retrieved from https://doi.org/10.48550/arXiv.2012.08548

5. Wang, J., Zhang, W., Zang, Y., Cao, Y., Pang, J., Gong, T., Chen, K., Liu, Z., Loy, C. C., & Lin, D. (2021). Seesaw Loss for long-tailed instance segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 9695-9704). Retrieved from https://doi.org/10.48550/arXiv.2008.10032

6. Yaman, B., Mahmud, T., & Liu, C.-H. (2023). Instance-aware repeat factor sampling for long-tailed object detection. Retrieved from https://doi.org/10.48550/arXiv.2305.08069

7. Menon, A. K., Jayasumana, S., Rawat, A. S., Jain, H., Veit, A., & Kumar, S. (2021). Long-tail learning via logit adjustment. International Conference on Learning Representations. Retrieved from https://doi.org/10.48550/arXiv.2007.07314

8. Ren, J., Yu, C., Sheng, S., Ma, X., Zhao, H., Yi, S., & Li, H. (2020). Balanced Meta-Softmax for long-tailed visual recognition. Advances in Neural Information Processing Systems. Retrieved from https://doi.org/10.48550/arXiv.2007.10740

9. Alexandridis, K. P., Elezi, I., Deng, J., Nguyen, A., & Luo, S. (2025). Fractal calibration for long-tailed object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Retrieved from https://doi.org/10.48550/arXiv.2410.11774

10. Tang, S., He, F., Huang, X., & Yang, J. (2019). Online PCB defect detector on a new PCB defect dataset. Retrieved from https://doi.org/10.48550/arXiv.1902.06197

11. Lv, S., Ouyang, B., Deng, Z., Liang, T., Jiang, S., Zhang, K., Chen, J., & Li, Z. (2024). A dataset for deep learning based detection of printed circuit board surface defect. Scientific Data, 11, Article 811. Retrieved from https://doi.org/10.1038/s41597-024-03656-8

12. Ultralytics. (2026). Model prediction with Ultralytics YOLO. Retrieved from https://docs.ultralytics.com/modes/predict/

A. Verlan, Y. Fastovets

LONG-TAIL BALANCING FOR IMPROVING RECALL OF RARE DEFECT CLASSES WITHOUT PRECISION DEGRADATION IN PCB DEFECT DETECTION

The paper considers the problem of long-tail balancing in automated optical inspection systems for printed circuit boards, Automated Optical Inspection (AOI) for Printed Circuit Board (PCB). This distribution is typical for manufacturing data: several defect categories occur frequently, whereas rare but technologically important classes are represented by a small number of annotated examples. As a result, a neural detector tends to learn frequent classes more reliably and may miss defects located in the tail of the distribution. In PCB quality control this problem is important because a missed real defect can lead to product failure, while an uncontrolled growth of false positive detections increases the workload of a human operator and reduces the throughput of the production line. The main idea of the study is not to increase recall at any cost, but to improve recall for rare defect classes while preserving an acceptable level of precision and a controlled number of False

Positive (FP) detections per board. The paper describes a methodological scheme for combining training-time and post-processing approaches: controlled sampling, hard negatives, Focal Loss, Class-Balanced Loss, Equalization Loss v2, Seesaw Loss, two-stage training, calibration, and class-specific threshold selection. Particular attention is paid to the evaluation protocol, which includes recall at a target precision level, FP per board, macro-recall, per-class precision-recall curves, and inference latency. The practical value of the work lies in a controlled scheme for tuning an AOI detector in which the improvement of rare-class detection quality is not considered separately from production constraints and the risk of precision degradation.

Keywords: AOI, PCB, printed circuit boards, long tail, class imbalance, neural detection, recall, precision.

Надійшла до редакції: 12.08.2026
Прийнята до друку: 02.09.2026
Опубліковано: 30.10.2026
© 2026 Верлань А. А., Фастовець Є. Р.
Цей матеріал ліцензовано за умовами CC BY 4.0.https://creativecommons.org/licenses/by/4.0/