비즈니스

웨이스는 VEXPLOR에서 AI가 자신 없는 불량 판정을 어떻게 사람에게 넘기는가

자신 있는 오답을 줄이는 3가지 장치

방동걸 (주)웨이스 대표이사
2026년 10월 7일

공장 AI가 가장 위험해지는 순간은 틀릴 때가 아니라 틀린 답을 자신 있게 낼 때입니다. 흔한 불량은 학습한 사례가 많아 잘 맞히지만, 몇 달에 한 번 나오는 불량은 사례가 적어 엉뚱한 답이 나오기 쉽습니다. 그 답이 확신에 찬 모습으로 나오면 현장은 그대로 따르기 쉽고, 오답 하나가 한 로트의 손실이 됩니다.

제조 AI 운영체제 VEXPLOR를 만드는 웨이스는 AI의 정확도를 올리는 것만으로는 이 위험을 막을 수 없다고 봅니다. 그래서 VEXPLOR에는 AI가 스스로 확신하지 못하는 판정을 사람에게 넘기고, 들어오는 데이터가 학습 때와 달라지면 감지하고, 판정마다 근거를 함께 남기는 장치를 둡니다. 이 글에서는 세 장치가 각각 무엇을 막는지, 그리고 아직 풀지 못한 것이 무엇인지 적습니다.

01

AI의 오답은 대개 학습한 사례가 적은 곳에서 나옵니다

AI 모델은 학습한 데이터와 비슷한 입력에서 잘 맞힙니다. 그런데 공장의 불량 데이터는 몇몇 유형에 몰려 있습니다. 스크래치나 찍힘처럼 매일 나오는 불량은 사진이 수천 장씩 쌓이고, 원자재 로트가 바뀔 때만 나오는 불량은 몇 장뿐입니다. 사례가 적은 유형을 만나면 모델은 비슷해 보이는 흔한 유형 쪽으로 답을 끌어당깁니다.

문제는 이때도 모델이 확신을 숫자로 내놓는다는 점입니다. 예를 들어 화면에는 「찍힘 92%」처럼 그럴듯한 값이 뜨는데, 그 값이 맞는지 틀린지는 화면만 봐서는 알 수 없습니다. 그래서 오답을 줄이는 일은 두 가지로 나뉩니다.

- **모델이 확신하지 못하는 판정을 걸러 내는 일** — 확신 값이 낮은 판정은 사람이 본다
- **모델이 확신해도 믿으면 안 되는 상황을 알아채는 일** — 입력이 학습 때와 달라졌는지 본다

웨이스가 VEXPLOR에 둔 장치는 이 두 가지에, 판단 근거를 남기는 장치 하나를 더한 것입니다.

장치	막는 것	사람이 하는 일
신뢰도 미달 판정 넘김	애매한 판정의 자동 처리	넘어온 판정 확인
데이터 드리프트 감지	바뀐 입력에서의 확신	바뀐 조건 확인
판단 근거 기록	근거 없는 판정 수용	근거와 현장 대조



VEXPLOR에 둔 3가지 장치 — 신뢰도 미달 판정 넘김, 데이터 드리프트 감지, 판단 근거 기록

02

VEXPLOR의 비전 AI는 확신이 기준에 못 미치는 판정을 사람에게 넘깁니다

VEXPLOR의 비전 AI는 구조가 서로 다른 여러 모델을 함께 써서 검사 이미지를 판정합니다. 이렇게 모은 판정의 확신 값이 기준보다 낮으면, 그 이미지는 자동으로 합격·불합격을 정하지 않고 사람에게 넘깁니다.

함께 쓰는 모델은 EfficientNet-B3, ResNet50, MobileNetV2 세 가지입니다. 여러 모델의 판정을 모으면, 일반적으로 구조가 다른 모델끼리 의견이 엇갈리는 이미지일수록 모은 확신 값이 낮게 나옵니다. VEXPLOR는 확신 값이 기준에 못 미치는 이미지를 사람에게 넘기도록 설계했습니다. 세 모델을 함께 쓰면서도 이미지 한 장의 판정을 0.1초 안에 끝내는 것을 목표로, 모델을 TensorRT INT8로 경량화하도록 설계했습니다.

이 장치가 바꾸는 것은 사람이 보는 이미지의 양과 종류입니다. 사람이 전수를 보던 검사가, AI가 애매하다고 판단한 것만 사람이 보는 검사로 바뀝니다. 드문 불량은 확신이 낮게 나오는 경우가 많아 이 장치로 사람에게

넘어올 가능성이 큽니다. 다만 확신이 높게 나오는 오답은 이 장치로 걸러지지 않으므로 다음 장치가 필요합니다.

기준을 높이면 사람에게 넘어오는 양이 늘고, 낮추면 AI가 혼자 정하는 양이 늘어납니다. 기준을 어디에 둘지는 놓쳐도 되는 불량이 얼마나 되는지에 달려 있어서, 같은 기준을 모든 공장에 그대로 쓸 수는 없습니다.

03

들어오는 데이터가 학습 때와 달라지면 VEXPLOR가 먼저 감지합니다

신뢰도 기준만으로는 부족할 때가 있습니다. 조명을 바꾸거나, 카메라 위치가 조금 움직이거나, 원자재 공급처가 바뀌면 이미지의 분포 자체가 달라집니다. 이때 모델은 처음 보는 입력에도 높은 확신을 낼 수 있고, 확신 값이 높으니 신뢰도 기준도 그대로 통과합니다.

그래서 VEXPLOR는 판정 하나하나가 아니라 들어오는 데이터 전체의 분포를 봅니다. 최근 입력의 분포가 학습 데이터의 분포에서 얼마나 멀어졌는지를 KL 발산(두 분포가 얼마나 다른지 나타내는 값)으로 계산하고, 일정 수준을 넘으면 드리프트로 판정합니다. 이 판정은 「모델이 틀렸다」가 아니라 「모델이 믿을 수 있는 범위 밖으로 나가고 있다」는 신호입니다.

드리프트가 감지되면 VEXPLOR는 모델이 크게 틀리기 전에 재학습을 시작합니다. 사람이 할 일은 현장에서 무엇이 바뀌었는지 확인하는 것입니다. 바뀐 것이 조명 교체처럼 의도한 변화인지, 설비 이상처럼 고쳐야 할 변화인지는 데이터 분포만으로는 알 수 없기 때문입니다.

04

VEXPLOR는 판정마다 근거를 함께 남겨 사람이 현장과 대조하게 합니다

세 번째 장치는 판정의 이유를 남기는 것입니다. VEXPLOR는 판정에 영향을 준 요인을 SHAP(각 입력이 결과에 얼마나 기여했는지 계산하는 방법)으로 계산해 보여 주고, 추론 과정을 로그로 남깁니다.

근거가 함께 보이면 사람은 판정을 통째로 받아들일지 무시할지만 놓고 고르지 않아도 됩니다. 「이 불량 판정은 표면 반사가 큰 영역 때문에 나왔다」는 근거가 보이면, 검사원은 그 영역이 실제 결함인지 조명 반사인지를 직접 확인할 수 있습니다.

05

VEXPLOR는 금속 표면의 반사를 판정 전에 촬영 단계에서 먼저 줄이도록 설계합니다

금속 부품 검사에서 대표적인 자신 있는 오답의 하나는 빛 반사를 결함으로 읽는 경우입니다. 난반사가 심한 금속 면에서는 반사광이 이미지에 찍힘이나 스크래치처럼 나타나기 때문입니다. 이미지에 반사가 남아 있으면 모델이 아무리 정확해도 이 오답은 줄지 않습니다.

그래서 VEXPLOR는 이 오답을 모델에 맡기기 전에 촬영 단계에서 먼저 줄이도록 설계했습니다. 빛을 흡수하는 암실 안에서 편광 카메라 여러 대로 찍고, 생성형 모델(GAN)로 두 단계에 걸쳐 반사를 걷어 낸 뒤 판정합니다. 목표는 난반사가 심한 금속 면에서도 0.1mm 결함을 찾는 것입니다. 촬영 단계에서 반사를 다 걷어 내지 못하면, 남은 반사는 아래 예시처럼 검사원이 판정 근거를 보고 결함과 구분합니다. 이 촬영 단계도 아직 현장 수치는 없습니다.

06

원자재 로트가 바뀐 날, 세 장치는 이런 순서로 움직입니다

세 장치는 따로 움직이지 않고, 한 사건 안에서 차례로 작동합니다. 아래는 이해를 돕기 위한 예시입니다.



원자재 로트가 바뀐 날 세 장치가 움직이는 순서를 그린 예시 — 입력 변화, 드리프트 감지, 애매한 판정 넘김, 근거 확인

1. **입력이 달라집니다** — 금속 판재의 공급 로트가 바뀌어 표면 광택이 조금 달라집니다. 사람 눈에는 같은 판재지만 카메라 이미지의 밝기 분포는 달라집니다

2. **드리프트가 먼저 감지됩니다** — 개별 판정이 틀리기 전에, 최근 이미지의 분포가 학습 데이터에서 멀어졌다는 것이 감지되고 재학습이 시작됩니다. 담당자는 이 시점에 로트 변경을 확인합니다
3. **애매한 판정이 사람에게 넘어옵니다** — 광택 차이 때문에 확신 값이 기준에 못 미치는 이미지가 늘고, 그 이미지들은 자동 처리되지 않고 검사원에게 넘어옵니다
4. **검사원이 근거를 보고 결함과 반사를 가립니다** — 검사원은 판정 근거로 표시된 영역이 실제 찍힘인지 광택 반사인지 확인합니다. 이렇게 확인한 판정은 새 로트 조건으로 재학습할 때 쓸 수 있습니다

장치가 하나만 있으면 이 사건의 어느 한 단계에서 오답이 그대로 흘러갑니다. 드리프트 감지만 있으면 재학습은 시작되지만 그사이의 판정은 자동으로 처리되고, 신뢰도 기준만 있으면 높은 확신으로 나온 오답이 걸러지지 않습니다. 근거 기록이 없으면 검사원은 넘어온 판정을 처음부터 다시 봐야 합니다.

07

VEXPLOR는 검사원의 판정 기준을 라벨링 체계로 옮겨 적는 방법을 함께 설계합니다

사람에게 넘어온 판정을 검사원이 확인하는 일은 한 번 보고 끝나지 않습니다. 숙련 검사원이 눈으로 결함과 반사를 구분하는 기준은 대개 그 사람의 경험에 담겨 있고, 글로 적혀 있지 않습니다. VEXPLOR는 이 판단 기준을 라벨링 체계로 옮겨 적는 방법론을 함께 설계하고 있습니다.

검사원이 확인한 판정이 이렇게 기록으로 남으면, 드문 불량의 학습 사례로 쓸 수 있습니다. 드문 불량의 정확도가 사례가 쌓여야 나아진다는 한계는 그대로입니다. 다만 쌓이는 사례에 숙련 검사원의 판단 기준이 함께 남을 수 있다는 점이 다릅니다.

08

웨이스가 아직 풀지 못한 것도 있습니다

세 장치는 오답이 현장에 그대로 흘러가는 것을 줄이지만, 드문 불량을 잘 맞히게 하는 장치는 아닙니다. 아직 남은 일은 아래와 같습니다.

- **드문 불량의 정확도 자체** — 사례가 적은 유형은 판정 사례가 쌓여야 나아집니다. 쌓이는 속도는 그 불량이 얼마나 자주 나오느냐에 달려 있습니다
- **기준값의 균형** — 사람에게 넘어오는 양과 놓치는 불량 사이의 균형은 공장의 검사 조건마다 다시 맞춰야 합니다
- **현장 수치** — VEXPLOR의 비전 AI는 정부 자율형공장 구축 지원사업 현장에서 구현·검증 중이라, 사람에게 넘어온 비율이나 놓친 불량 수 같은 수치는 아직 집계 전입니다



이 글을 한 장으로 요약하면 이렇습니다 — 웨이스는 VEXPLOR에서 자신 없는 판정은 사람에게 넘기고, 바뀐 입력은 감지하고, 판정마다 근거를 남깁니다

공장 AI의 오답을 0으로 만들 수는 없으니, VEXPLOR는 AI가 자신 없는 판정과 믿을 수 없는 상황을 사람이 먼저 보게 합니다.

WACE(웨이스)는 공장이 스스로 판단하고 멈추지 않게 하는 제조 AI 운영체제 VEXPLOR를 만드는 회사다.

09

함께 읽어보세요

- 웨이스는 공장장이 AI 판단을 승인하고 그 범위를 정할 수 있도록 VEXPLOR에서 AI의 권한을 단계로 나눕니다
- 제조업 승계가 막히는 문제를 웨이스는 VEXPLOR로 어떻게 풀려 하는가 — 판단 근거를 묶고 권한을 단계로 넘기는 방법
- 상관은 알람을 만들고, 인과는 결정을 만든다
- AI에게 설비를 맡기려면, 틀린 명령이 나가기 전에 멈출 수 있어야 합니다

원문 blog.vexplor.com/post/how-vexplor-routes-unsure-ai-calls-to-people
발행 2026년 10월 7일 · 방동걸 (주)웨이스 대표이사

회사 주식회사 웨이스 (WACE Inc.) · 대표이사 방동걸 · 사업자등록번호 439-87-02852
주소 (18462) 경기 화성시 동탄구 동탄영천로 150, 에이동(A동) 1702-1708호 (영천동, 실리콘앨리동탄)
연락처 전화 070-4730-0629 · contact@wace.me
홈페이지 wace.me · linkedin.com/company/wace-inc

원문 홈페이지 링크드인