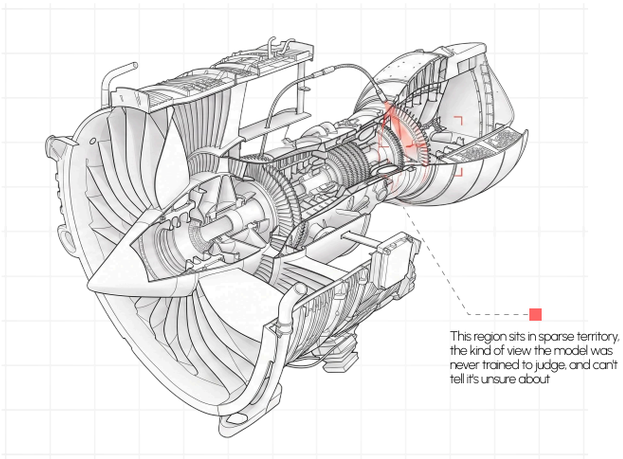
Case Study - Aerospace

WHY AN AUTOMATED JET ENGINE INSPECTION MODEL MISSES THE CRACKS IT MOST NEEDS TO CATCH

An aerospace illustration of how Squint Vision Studio maps the blade images a borescope inspection model can judge reliably, and what to do with the ones it cannot.



A jet engine is inspected from the inside without being taken apart. An inspector threads a borescope, a flexible optical probe with a camera on the end, through an access port and works along the compressor and turbine stages one blade at a time.

The light is poor. The blades number in the hundreds and look nearly identical. The inspector is searching for damage measured in fractions of a millimeter, on a curved metal surface, through a lens the width of a pencil.

The industry has been automating this work, and the case for it is obvious. Deep learning models now scan borescope footage and flag damaged blades for an engineer to review, reporting accuracy in the high eighties.

That number sounds like a system ready to be trusted. What it leaves out is the subject of this case study.

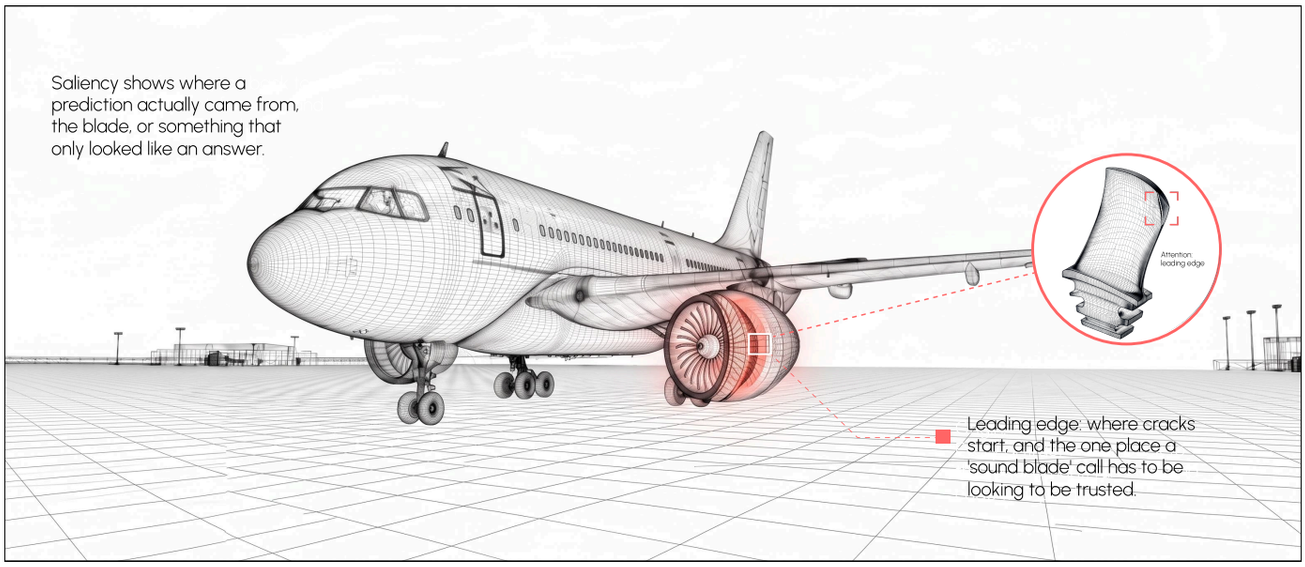
THE ONE DEFECT THAT ENDS ENGINES

Blade damage is not all the same, and the differences matter more than the totals.

Erosion, worn coating, nicks and dents are common. They are visible, they are well understood, and an engine can usually keep running with them under defined limits.

A crack is a different category of problem. In a blade spinning at thousands of revolutions per minute, a crack grows under load until the blade separates, and a liberated blade destroys the engine around it. Cracks are also where the smaller defects lead, since a nick or a patch of corrosion is often where the first one starts.

So the defect that matters most is the one an inspection exists to find, and it is also the one that is hardest to see.



WHY A CRACK HIDES IN PLAIN SIGHT

A crack is often less than a millimeter across. On a borescope image it appears as a thin pale line against textured metal, and whether it appears at all depends on the angle of view.

That angle dependence is the part that makes automation difficult. The same crack can be obvious in one frame and invisible in the next, because a change of a few degrees alters how light falls into it. Published work puts the manual miss rate at roughly 23 percent for cracks under 0.3 millimeters, and that figure describes trained human inspectors.

Training data compounds the problem from the other direction. Engines are built specifically not to crack, so cracks are rare in any real dataset. A model will see thousands of examples of erosion and a handful of cracks.

The result is an inspection system that is strongest on the damage that matters least and weakest on the damage that ends engines. An accuracy figure hides this, because it is an average, and the average is dominated by the common defects the model handles well.

WHAT A CONFIDENCE SCORE ACTUALLY MEASURES

You might expect the model to hesitate on a crack it cannot handle. It has no way of doing so. A faint crack seen from an awkward angle falls in a part of the model's experience where it has seen almost nothing. The model matches it to the closest thing it knows, and for a thin line on textured metal, the closest thing it knows is usually an undamaged blade.

It then reports high confidence, because confidence measures how strongly the model prefers one answer over the others. It says nothing about whether the model was equipped to judge that image at all.

A confident pass on a healthy blade and a confident pass on a cracked one are identical coming out.

TWO QUESTIONS WORTH ASKING BEFORE DEPLOYMENT

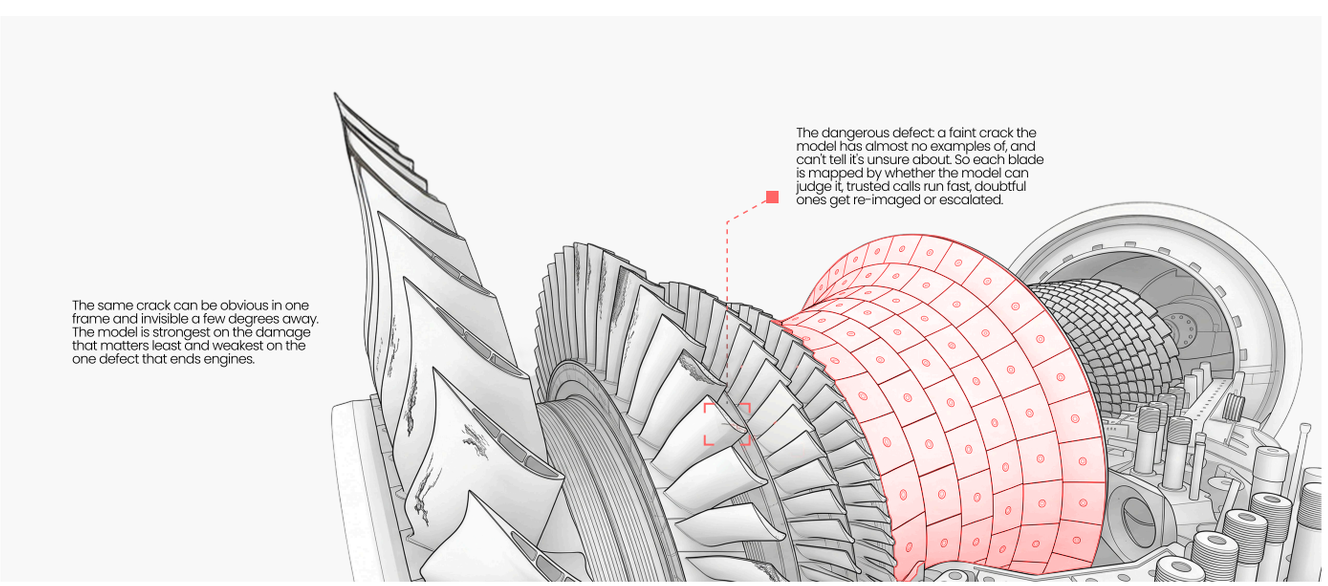
Squint Vision Studio approaches this through the model's own internal state rather than its output, and for a borescope classifier there are two separate things worth establishing.

The first is where the model has experience. Vision Studio extracts the internal representations the network builds as it processes each image, then clusters them to produce a view of how the training data is organized inside the model. Common defects form dense, well-populated regions. Small cracks and awkwardly angled ones sit near the boundaries, in sparse territory where the model's separation between damaged and sound metal breaks down. Regions whose accuracy falls below a chosen threshold are marked ambiguous.

The second question is different, and it matters for a defect this small. Saliency analysis shows which parts of an image drove a given prediction. Squint's published work uses this to catch a model that reaches the right answer through the wrong features, the example being a network that predicts airplane because the sky is prominent rather than because of the aircraft.

The borescope equivalent is a model that calls a blade sound while attending to the platform or the blade root rather than the leading edge where cracks initiate. That model will score well on a test set and fail on the first crack that appears somewhere it was never really looking. The map tells you where the model has data. Saliency tells you whether it used the right part of it.

An inspection team can establish both before the model is trusted on real engines, rather than discovering the gap after a blade fails in service.



THREE PATHS THROUGH AN INSPECTION

Vision Studio exports the analysis as a watchdog, a lightweight decision layer that runs alongside the model rather than replacing it. Its job is to decide when the base model can be trusted and when further processing is worth the cost. Each image takes one of three paths.

An image in a trusted region takes the fast path. The result is accepted and the sweep continues at full speed, which covers the routine erosion and coating wear that make up most of any inspection.

An image in an ambiguous region is not accepted at face value. Squint's published pipeline escalates this path to a specialist model, and the reasoning behind that is worth stating. The general classifier is separating many defect classes at once. Inside the ambiguous region the problem is narrower, usually faint crack against textured sound metal, and a model trained on that subset alone is solving an easier problem than the one the general model faces. Squint reports measurable accuracy gains from exactly this arrangement in its published results.

A borescope inspection also offers something most deployments do not, because the probe is still inside the engine. Re-imaging the same blade from a different position costs seconds, and given how sharply crack visibility depends on angle, a view that settles the question may be a few degrees away. Whether to add that step is a design decision for the operator, but the physics makes it an unusually cheap one.

An image unlike anything in the training data takes the review path. It is retained for analysis rather than acted on, because it points at a condition the model was never trained to resolve.

There is a second signal available across a full sweep. When ambiguous calls cluster on one blade, one stage, or one lighting condition, the pattern says something a single frame does not. It suggests a region of the engine the model is systematically struggling with, and those cases can be grouped for closer examination together.

WHAT THE MAP CANNOT SEE

Mapping the model's experience catches cracks that fall in sparse or ambiguous regions, which is where the dangerous misses concentrate. It does not guarantee catching a crack that looks, to the model, like ordinary sound metal. If a defect sits squarely inside a region the model handles confidently, a map built around where the model lacks support will not necessarily flag it.

The map is also only as good as the data behind it. If the training set underrepresents a crack morphology the engine will actually produce, the region covering it may look better supported than it is.

And none of this makes the model better at finding cracks. It makes the model's limits visible, which is a different thing and the reason retraining alone does not solve the problem. Cracks vary in length, direction, branching and viewing angle, so new examples help only where they happen to land. The next unusual crack from an unusual angle will still be unfamiliar.

WHAT CHANGES ON THE LINE

An inspection built on a single accuracy number is trusting that figure to promise something it cannot, which is that the model is reliable on the one defect it most needs to find. The figure says the model is usually right. It does not say the model is right about cracks, and on cracks it is reliably worst.

Seeing where the model's experience ends changes what an operator can depend on. The model keeps its speed and consistency across the routine majority of blades. The calls it cannot support are separated out, given another look, grouped when they cluster, and passed to an engineer when nothing else settles them.

For a component spinning inside a running engine, that is the difference between an inspection that reports a number and one that knows which of its own answers to doubt.

This case study describes a representative scenario. The defect types, model architectures, detection difficulty and the documented manual miss rate for cracks under 0.3 millimeters are drawn from published research on automated borescope inspection. Squint Vision Studio's capabilities, including embedding extraction, clustering-based region mapping and exportable watchdog modules, are described as documented by Squint Cognition rather than measured from this scenario.