Sim-to-Real Transfer for Document-Style Geoscience Images

The standard way to build a document-image model is to scan documents, label them, and learn from the labels. Raster well-log digitisation cannot follow that recipe, because nobody ever labelled which pixels in a scanned log belong to which curve, and labelling enough of them by hand is the exact manual cost the project exists to remove. Our alternative was to train end to end on a corpus that contains no real document at all: 20,000 procedurally generated logs from which a final version-two training set of 15,000 curves was drawn, every pixel mask manufactured by construction rather than annotated, and the model never shown a single scanned page during training. The wager in that choice is sim-to-real transfer, the question of whether a network raised in a generator survives the distribution shift to real scans. We argue it does, and we argue why. Evaluated on reconstructed curves against the native Log ASCII Standard data of a public Texas archive of 136,771 scanned TIF documents and 7,781 LAS files, the same fully synthetic training run reaches a peak coefficient of determination of 0.9891, with the model encountering those real scans for the first time at inference. The contribution of this document is not the headline number but the design discipline behind it: which axes of the synthetic corpus had to bracket the real archive's range for transfer to hold, why breadth of randomisation rather than photorealism is what carries the jump, where the reality gap reopens on degraded scans, and how we measured all of it so the claim that transfer survived is auditable rather than asserted. We present three interactive readers built directly on the engagement figures, walk the corpus-design decisions axis by axis, and state the operating rule we now apply to any document-style transfer problem of this shape.

The EarthScan Team
Research, engineering, and field

VEERNET / SIM-TO-REAL / TRANSFER-LEARNING / SYNTHETIC-DATA / DOMAIN-RANDOMISATION / DOCUMENT-IMAGE-ANALYSIS

CONTENTS

01 What we built and what it cost the assumption book

02 Why the conventional pipeline is unavailable, not merely inconvenient

03 Why document-style images make the bet more reasonable than it sounds

04 The mechanism we relied on, stated plainly

05 The corpus design axes that decided transfer

06 The jump held, and it held for an auditable reason

07 Grading on the deliverable, not the synthetic mask

08 Why the generator's freedoms are also its honesty controls

09 What a survived jump licenses and what it does not

10 The operating rule we now carry into document-style transfer

11 Limitations

12 References

13 Get the full whitepaper

> "
>
> The model had never seen a real scanned log when it was asked to digitise 136,771 of them. Every page it learned from was drawn, not photographed. The whole question is whether that gap is survivable, and the archive answered it.
>
> "

I

EXECUTIVE SUMMARY

A model raised in a generator, deployed on real paper

What we built and what it cost the assumption book

We built VeerNet, our raster well-log digitiser, the way almost nobody builds a document-image model: on a training set that contains no real document. The corpus was procedural end to end. A generator drew 20,000 synthetic logs, each one rendered with the curve traces, the grid lines, the depth ticks and the printed track labels a real scanned log carries, and because the generator drew every pixel it also knew every pixel, so the segmentation mask came for free with no human in the loop. The final version-two training set drew 15,000 of those curves. At no point in training did the network see a scanned page.

The reason to take that risk is structural. The raw material for this problem is decades of scanned paper logs sitting in operator and regulator archives, and none of them carries a label that says which pixels are the gamma-ray trace and which are the grid line behind it. A segmentation network needs that label for thousands of images, and producing it by hand is precisely the manual bottleneck the whole effort is meant to remove. You cannot annotate your way out of an annotation problem, so we manufactured the data instead and bet on sim-to-real transfer to carry the model across the gap to real scans.

The bet paid. Evaluated on reconstructed curves against the native LAS data of a public Texas archive that holds 136,771 scanned TIF documents and 7,781 LAS files, the fully synthetic training run reached a peak coefficient of determination of **0.9891**, with the model meeting those real scans for the first time at inference. This document is not a victory lap on that number. It is the engineering account of why the jump held, told through the specific synthetic-corpus design choices that made real TIFs look, to the model, like one more draw from the training distribution, and the choices that would have quietly broken it.

THE TRANSFER, IN FOUR NUMBERS

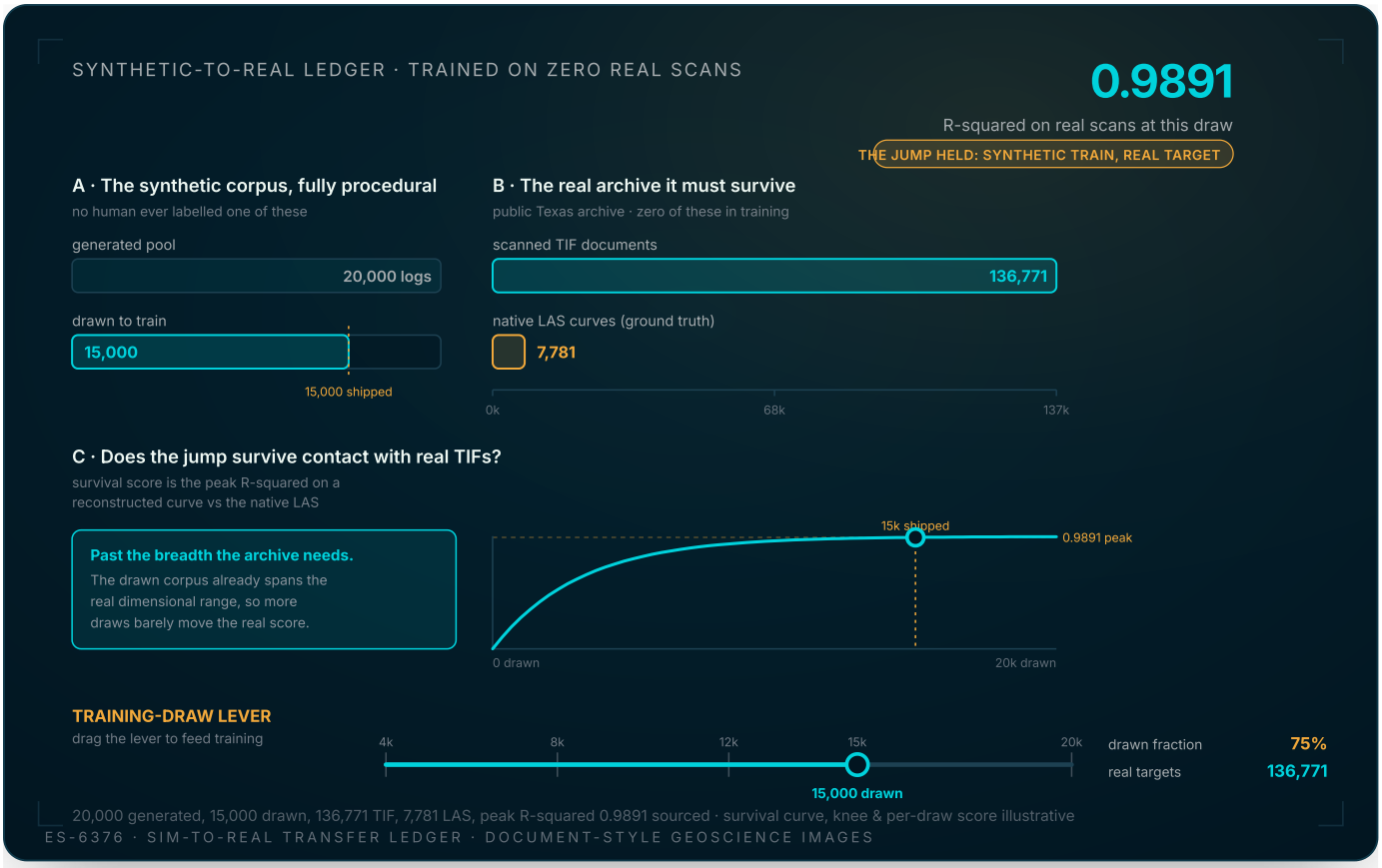
20,000
Synthetic logs generated, zero real

15,000
Curves drawn into the final training set

136,771
never in training
Real scanned TIFs the model digitises

0.9891
from synthetic-only training
Peak R-squared on real scans vs LAS

The first reader on this page is the ledger that holds those numbers together: the synthetic corpus on one side, the real archive on the other, and the survival score that says whether the model that learned from the first can serve the second.



One reader over the synthetic-to-real transfer ledger for a model trained on zero real scans. Exhibit A is the procedural corpus: 20,000 synthetic logs were generated and the shipped version-two training set drew 15,000 of them, none labelled by a human. Exhibit B is the real target the model must survive: 136,771 scanned TIF documents and 7,781 native LAS curves in the public Texas archive, none of which were ever in training. Exhibit C is the transfer verdict: the peak coefficient of determination on a reconstructed curve, measured against the native LAS, reached 0.9891 from that fully synthetic training set, and the lever shows how the real-scan score saturates as more of the generated pool is drawn in. The 20,000 generated, 15,000 drawn, 136,771 TIF, 7,781 LAS, and 0.9891 peak are sourced from the engagement archive; the transfer-survival curve, its diminishing-return knee, and the per-draw interpolated score are illustrative.

Read the ledger left to right and the argument of the whole whitepaper is already visible. The generated pool and the drawn training set are both synthetic. The TIF and LAS counts are both real and were both absent from training. The bridge between them is a single number, and that number held at the draw we shipped. Everything that follows is the explanation of why a bridge that looks this unlikely on paper carried real traffic.

II

THE PROBLEM AND WHY NOW

The recipe document-image models use does not apply here

Why the conventional pipeline is unavailable, not merely inconvenient

The default pipeline for a document-image task is well worn: collect documents, label the regions of interest, train a network to reproduce the labels, and evaluate on a held-out slice of the same labelled distribution. It works because the training and test data are drawn from one source, so there is no distribution shift to survive. The model is graded on data that looks exactly like what it learned from.

Raster well-log digitisation breaks that pipeline at the first step. There is no labelled corpus of scanned logs and no realistic prospect of one. The archive is decades of paper at every scale, vintage and scan condition, and the labels a segmentation model needs, a per-pixel assignment of curve identity, were never produced and would cost more to produce at scale than the digitisation they are meant to enable. The cruelty of it is that the labelling cost and the digitisation cost are the same cost: a human who can label which pixels are the gamma-ray trace has, in the act of labelling, very nearly done the digitisation by hand. A pipeline that needs that human first has not removed the bottleneck. It has renamed it.

So the synthetic route is not a shortcut taken to save effort on a problem that had an easy labelled answer. It is the only route that changes the economics, because it replaces a per-image human cost with a one-time generator-design cost. The prize that justifies the attempt at archive scale is what the digitised curves unlock downstream, where decades of legacy logs become queryable inputs to the upstream workflows that depend on them [7]. That reframing is what makes sim-to-real the load-bearing question. The moment you decide to train on manufactured data, you have accepted a distribution shift between training and deployment as the price of admission, and the entire viability of the approach turns on whether that shift is survivable for this particular class of image.

Why document-style images make the bet more reasonable than it sounds

Training a model on rendered images and deploying it on real ones is not new, and the precedent matters because it tells you which problems the trick works on. Scene-text recognition was solved at scale by training on rendered text and reading real photographs, with no real labelled text in the loop [5]. The reason it worked there is the reason it has a chance here: a printed document is a far more constrained visual object than a natural scene. A scanned well log is monochrome, axis-aligned, built from a small alphabet of printed primitives, curve traces, grid lines, ticks and labels, and corrupted by a bounded set of scan artefacts. The space of real scans is wide, but it is not the open-ended chaos of natural imagery. It is a space a generator can be made to span.

That is the precondition the rest of this document leans on. Sim-to-real transfer is not a universal solvent; it works when the generator can be dialled to cover the real distribution, and document-style geoscience images are a case where covering the distribution is achievable because the distribution, for all its breadth, is describable. The job then becomes a design job: identify the axes along which real scans vary, bound each one with a real archive statistic, and randomise across all of them widely enough that the real archive becomes an unremarkable sample of the synthetic one [1].

III

OUR APPROACH

Breadth on every axis, not realism on any one

The mechanism we relied on, stated plainly

The principle we built on is domain randomisation: if you randomise a generator's nuisance parameters widely enough, the model stops treating any particular setting of them as signal and learns only the invariant it actually needs, which here is the shape of a curve trace against printed clutter [1]. The counterintuitive evidence behind it is that deliberately unrealistic but widely varied synthetic data can beat carefully photorealistic synthetic data on the real target, because photorealism narrows the distribution toward one convincing point while randomisation spreads it across the whole space the real data might occupy [2]. We took that evidence seriously and spent our design effort on breadth rather than on making any single render look like a convincing scan.

There was a road we deliberately did not take. The adversarial branch of transfer, unsupervised domain adaptation, aligns the feature distributions of source and target by pulling them together during training [4]. It is powerful, but it needs real target data in the loop and a training procedure that is harder to keep stable, and it adapts to the specific real distribution you trained the alignment against. Randomisation needs no real data at training time and produces a model that is robust to a band of conditions rather than tuned to one. For an archive as heterogeneous as decades of scanned paper, robustness to a band was the property we wanted, so we widened the source instead of aligning to the target.

The corpus design axes that decided transfer

The discipline that turns this principle into a working corpus is that every randomisation axis is bounded by a real archive statistic, never by what was convenient to render. The synthetic span on each axis has to bracket the real span on that axis, on both ends, or a slice of the real distribution falls outside everything the model ever saw, and that slice is

exactly where transfer breaks first. The second reader makes this concrete: it lays each design axis's synthetic span over the real archive's span and lets you narrow the generator until you can watch a part of the real distribution fall out of coverage.

Image width

- Real archive spans 3,200 to 12,800 pixels wide on a scanned log
- Generator dialled to bracket the full range, not a comfortable middle
- The widest scans are where an under-ranged corpus loses transfer first
- Bounded by an archive statistic, not by what was convenient to render

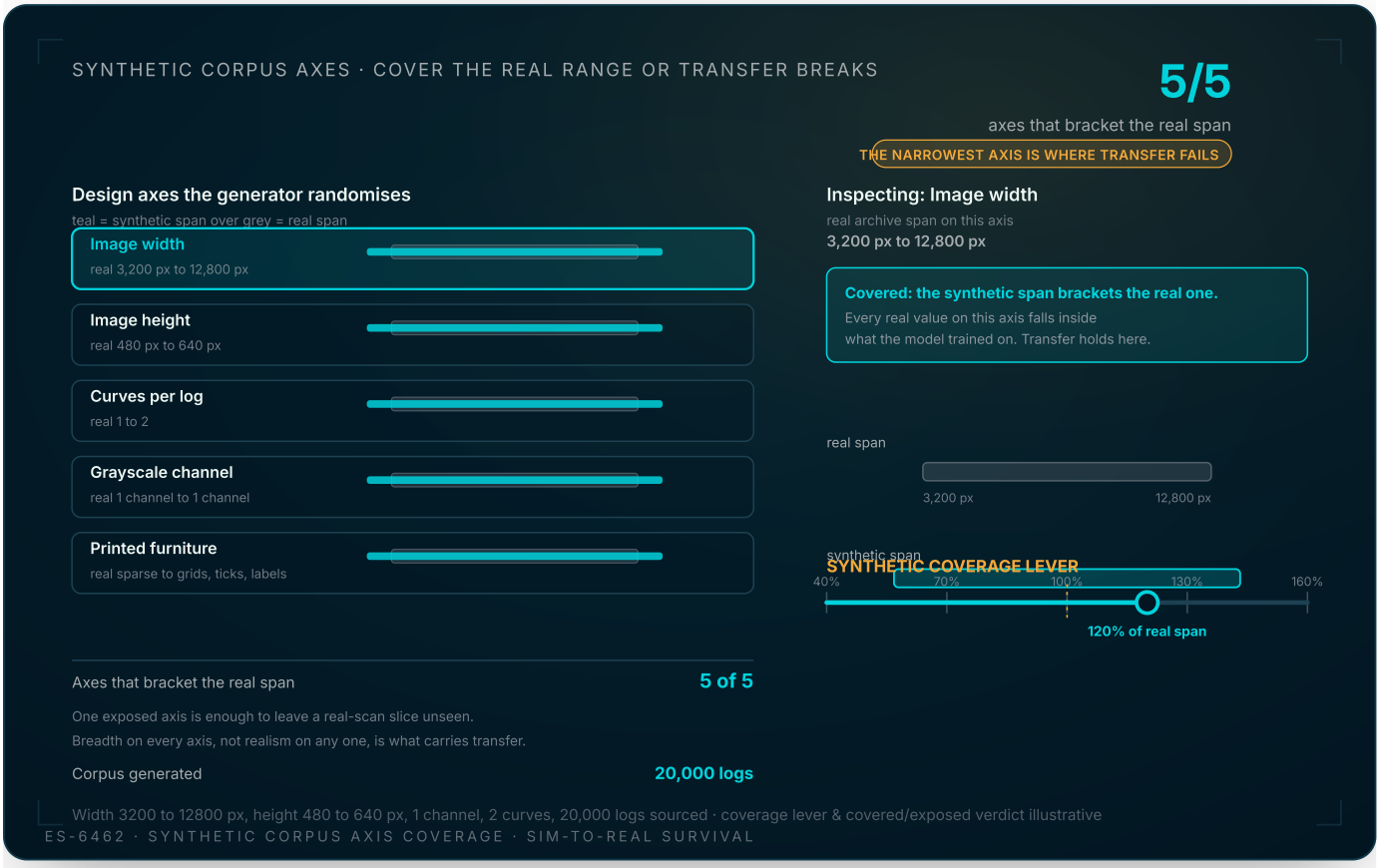
Image height

- Real archive spans 480 to 640 pixels of usable log height
- Held variable per render so the model never assumes a fixed page
- Narrow on its own, but coupled to width it sets the aspect-ratio spread
- A fixed height would have taught a shortcut the real scans punish

Curves and channel

- One to two curves per log, a single grayscale input channel
- Matches how the real scans arrive: monochrome, few traces per track
- Two curves is the multiclass setting the final corpus was built for

- Colour would have been realism the real archive does not actually carry



A coverage reader over the synthetic corpus design axes that decided whether transfer to real scans survived. Each axis shows the procedural generator's synthetic span laid over the real archive's span: image width 3,200 to 12,800 px, image height 480 to 640 px, one to two curves per log, a single grayscale channel, and the printed furniture a scanned log carries. Click an axis to inspect it, then drag the coverage lever to set how wide the generator was dialled on that axis as a fraction of the real span. The verdict reads covered when the synthetic span brackets the real one on both ends and exposed when a slice of the real distribution falls outside training, which is where sim-to-real transfer breaks first. The width, height, channel, curve count, and the 20,000-log corpus size are sourced from the engagement archive; the coverage-fraction lever and the covered or exposed verdict are illustrative aids for reasoning about the design.

The axes are not equally forgiving. Width is the dangerous one. Real scans run from 3,200 to 12,800 pixels wide, a factor of four, and the widest scans are visually the most different from a comfortable median render. A generator that drew only middling widths would look fine on synthetic validation and then meet a 12,800-pixel scan at inference that sits outside its entire training experience. Height, from 480 to 640 pixels, is narrow on its own, but coupled to width it sets the aspect-ratio spread, and a fixed height would have let the model learn a page-shape shortcut that real scans punish. The curve count, one to two per

log, and the single grayscale channel match how the real scans actually arrive: monochrome, few traces per track. Adding colour would have been realism the real archive does not carry, which is the trap photorealism sets, spending fidelity on a dimension the target does not vary along while leaving the dimensions it does vary along under-covered.

> **THE CORPUS-DESIGN RULE WE ADOPTED**
>
> Bound every randomisation axis by a measured statistic of the real archive, then dial the synthetic span to bracket the real span on both ends. Spend the design budget on covering the axes the real data varies along, especially the wide ones, not on photorealism along axes it does not. A generator that brackets every axis transfers; a generator that is photoreal on one axis and narrow on another does not.

IV

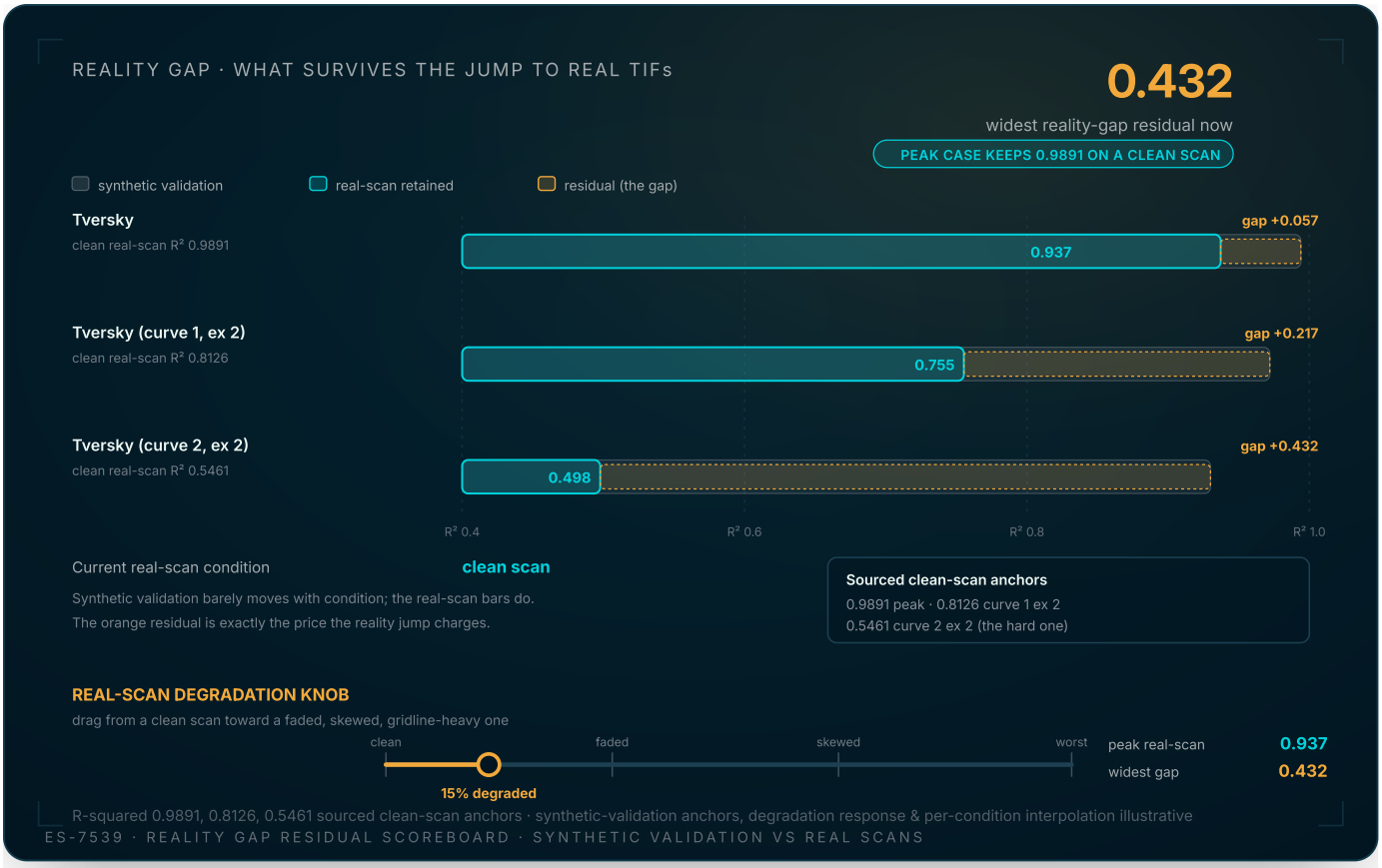
THE RESULT AND THE REALITY GAP

What survived, and where the gap reopens

The jump held, and it held for an auditable reason

The fully synthetic training run reached a peak coefficient of determination of 0.9891 on a reconstructed curve measured against the native LAS of the real archive. That is the headline, and on its own it is just an assertion. What makes it a finding rather than a press line is that it is reproducible from the corpus design: the model scored that well on a real scan because the real scan fell inside the synthetic distribution on every axis that mattered, which is the thing the coverage discipline of the previous section was engineered to guarantee. Transfer is not luck here. It is the visible consequence of bracketing the archive.

But a single peak number can hide a model that is fragile, so the honest version of the result is not the peak alone. It is the peak together with where the model loses ground, and a model trained only on synthetic data has a characteristic failure shape: it can ace its own synthetic validation set and still surrender accuracy on degraded real scans, because the synthetic validation set is drawn from the same generator the model learned from, while the degraded real scan is not. The distance between those two scores is the reality gap, and the third reader measures it directly.



A scoreboard for the reality gap: the distance between a model's score on held-out synthetic validation and its score on a real scanned document. Each row is a Tversky-trained reconstruction with its synthetic-validation R-squared in grey and the real-scan R-squared it retains in teal; the orange residual is the ground the reality jump costs. The clean-scan anchors are sourced from the engagement archive: a peak coefficient of determination of 0.9891 on the hardest curve-1 example, 0.8126 on a second curve-1 example, and 0.5461 on the harder second curve. Drag the degradation knob from a clean scan toward a faded, skewed, gridline-heavy one and the real-scan bars slide down while synthetic validation stays put, opening the residual. The three R-squared endpoints are sourced; the synthetic-validation anchors, the degradation response, and the per-condition interpolation are illustrative schematics.

The scoreboard reads the same fully synthetic training run three ways, against the sourced clean-scan anchors of 0.9891 on the hardest curve-1 example, 0.8126 on a second curve-1 example, and 0.5461 on the harder second curve. On a clean scan the model keeps almost all of its synthetic-validation score, which is what successful transfer looks like. Drag the degradation knob toward a faded, skewed, gridline-heavy scan and the real-scan bars slide down while synthetic validation stays put, opening the orange residual. That residual is the price the reality jump charges, and the fact that it is small on clean scans and grows on degraded ones is the most useful diagnostic the whole exercise produced, because it tells you exactly which real conditions to widen the generator against next.

THE SAME SYNTHETIC RUN, READ AGAINST REAL LAS

0.9891
Peak R-squared, hardest curve-1 example

0.8126
R-squared, second curve-1 example

0.5461
R-squared, harder curve-2 example

15,000
Synthetic curves that produced all three

The spread across those three numbers is itself a result. The second curve is genuinely harder than the first, and the model is honest about that, scoring 0.5461 where it scored 0.9891 on the easier trace. A model that had merely memorised its generator would not show a clean difficulty gradient on real data; the gradient is evidence that the model learned the geoscience, the shape of a curve against clutter, rather than the generator's habits. That distinction, learning the invariant rather than the source, is the whole point of randomisation, and the per-curve spread is what it looks like when it works.

> Transfer holding is not the peak number alone. It is the peak on a clean scan, the honest difficulty gradient across curves, and a reality-gap residual that stays small until the scan itself degrades. All three are deliverable-space facts, and all three checked out.

V

METHODS DEEP-DIVE

How we measured a claim about data we manufactured

Grading on the deliverable, not the synthetic mask

A claim that synthetic-only training transferred to real scans is only as good as the thing it is measured on, and the measurement has to be on real data the model never saw. Two methodological choices make the claim auditable rather than self-confirming.

The first is that we grade the deliverable, not the mask. A digitiser does not ship a segmentation mask; it ships a reconstructed curve, a value at every sampled depth, exported as a CSV that a petrophysics package reads. The transfer claim is therefore measured on the reconstructed curve against the native LAS ground truth, resampled to a common depth axis, using the coefficient of determination, mean absolute error and mean squared error. The LAS files in the public archive are real digital curves produced independently of any model, so grading against them is grading against ground truth that owes nothing to our generator. This is the cleanest possible test of transfer: synthetic in, real out, scored against a real reference.

The second is that the synthetic validation set and the real test set are kept rigorously separate, because conflating them is the classic way a synthetic-training result lies to its authors. Synthetic validation tells you the model learned the generator. Only real-scan evaluation tells you the model learned the world. We report both, and the reality-gap residual between them, precisely so that the difference is on the page rather than hidden in an aggregate. A whitepaper that reported synthetic validation alone and called it transfer would be measuring the wrong thing with confidence.

Why the generator's freedoms are also its honesty controls

Manufacturing the data creates a temptation that has to be actively resisted: the generator could be tuned, consciously or not, to produce exactly the scans the model finds easy, and the result would be a high number on a corpus the author quietly shaped to be flattering. The defence against that is built into the randomisation discipline. The generator's nuisance axes are bounded by archive statistics rather than by results, so the corpus cannot be narrowed toward the model's comfort zone without violating the coverage rule that the second reader visualises. A generator that brackets the real archive on every axis is a generator that cannot be secretly easy, because the real archive's hard cases are inside its range by construction.

The architecture is upstream context for this argument rather than its subject. VeerNet is an encoder-decoder segmentation network in the U-Net lineage [3], consuming the synthetic grayscale image and emitting the per-pixel mask the generator manufactures for free, with the sparse-foreground imbalance of a one-pixel curve trace handled by a Tversky objective whose precision and recall tilt is tunable [6]. None of that is novel, and none of it is what carried transfer. What carried transfer was the data, and the methods that matter for this whitepaper are the ones that let us measure transfer honestly on real scans rather than the ones that produced the network.

> **“Synthetic validation tells you the model learned the generator. Only real-scan evaluation tells you it learned the world. We reported both so the difference between them is on the page.”**
>
> — FROM OUR OWN EVALUATION PROTOCOL

VI

IMPLICATIONS AND ROADMAP

Where this transfers, and where it stops

What a survived jump licenses and what it does not

The result licenses a specific and valuable claim: for document-style geoscience images, a model can be trained at archive scale with no human labelling at all, because the synthetic corpus can be made to bracket the real distribution and the transfer survives. That is the claim that makes a 136,771-document archive tractable, and it is the claim the rest of the digitisation product is built on. It does not license the broader claim that synthetic-only training is always safe, and the reality-gap residual is the reason for the caution. Transfer held on clean scans and degraded gracefully on harder ones precisely because the generator's coverage was matched to the archive's variation; a different archive with variation the generator did not anticipate would reopen the gap.

The roadmap therefore runs through the residual. The reality-gap scoreboard does not just diagnose; it prioritises. Each real-scan condition where the residual grows, heavy gridlines, severe skew, faded ink, is a randomisation axis the generator should be widened along, and each widening is a one-time design cost that buys robustness across the whole archive rather than a per-image labelling cost. The work after this is not collecting and labelling real scans. It is reading the residual, finding the real condition that opened it, and dialling the generator to bracket that condition too. The corpus design is the product surface that improves, and it improves without ever asking a human to label a pixel.

The operating rule we now carry into document-style transfer

Three things settled into a rule we apply to any problem of this shape. Measure transfer on the deliverable against a real reference the model never saw, never on the synthetic validation set, because only the first is a test of the world rather than of the generator. Spend the corpus-design budget on bracketing every axis the real data varies along,

especially the wide ones, rather than on photorealism along axes it does not, because breadth carries the jump and fidelity to a single point does not. And keep the reality-gap residual visible, condition by condition, because the residual is simultaneously the proof that transfer held on the easy cases and the map of exactly where to widen the generator next. A document-image model raised entirely in a generator can serve a real archive of 136,771 scans, and it will do so for as long as the generator keeps bracketing the archive it is asked to read.

WHAT TO CARRY OUT OF THIS

1. A digitiser trained on **20,000** generated logs, drawing **15,000** into training and seeing **zero** real scans, reached peak R-squared **0.9891** on reconstructed curves against the native LAS of a real archive of **136,771** TIFs and **7,781** LAS files. Sim-to-real transfer survived end to end.
2. The conventional label-then-train pipeline is unavailable here, not just inconvenient: labelling a scanned log by hand is nearly the digitisation itself, so a pipeline that needs labels first has only renamed the bottleneck.
3. Transfer held because the synthetic corpus was dialled to bracket the real archive on every axis that matters, width 3,200 to 12,800 px, height 480 to 640 px, one to two grayscale curves. The narrowest axis is where transfer breaks first, so breadth on all of them beats realism on any one.
4. We graded the claim on the deliverable, reconstructed curves against real LAS the model never saw, and kept synthetic validation separate from real-scan evaluation so the reality-gap residual between them is on the page rather than hidden in an aggregate.
5. The reality gap stays small on clean scans and grows on degraded ones, which makes it a roadmap: each condition that reopens the gap is the next randomisation axis to widen, a one-time design cost rather than a per-image labelling cost.

GLOSSARY

Corpus draw

The number of synthetic logs actually pulled from the generated pool into the training set. The generator produced 20,000; the shipped version-two training set drew 15,000.

Domain randomisation

Randomising a generator's nuisance parameters, width, height, noise, furniture, so widely that the real domain becomes, to the model, just another draw from the synthetic distribution. The mechanism this whitepaper credits for the transfer.

LAS

Log ASCII Standard, the canonical text format for digital well-log curves. The ground truth a reconstructed curve is graded against and the format the deliverable is exported to.

R-squared

Coefficient of determination, the fraction of variance in the ground-truth curve explained by the predicted curve. The deliverable-space headline for a digitiser. Peak here, on real scans from fully synthetic training, is 0.9891.

Reality gap

The performance distance between a model's score on held-out synthetic data and its score on real data. A fully synthetic run can ace synthetic validation and still lose ground on real scans; the gap is what that loss is called.

Reality gap residual

The numeric size of the reality gap on a given example, synthetic-validation score minus real-scan score. The orange band in the third reader is exactly this quantity.

Sim-to-real transfer

The ability of a model trained in a simulator or generator to perform on real data it never saw in training. Here, a model trained only on procedurally drawn logs that must work on real scanned TIFs.

TIF

The raster image format the scanned paper logs arrive in. The 136,771 TIF files in the public Texas archive are the real documents the synthetic-trained model is asked to digitise.

Limitations

The strongest result, peak R-squared 0.9891, is a per-curve, per-example peak on the easier of the two curves, and the 0.5461 we report on the harder second curve is the honest other end of that same run; the transfer claim should be read across the full spread, not from the peak alone. The reality-gap residual and the degradation response shown in the third reader are illustrative reasoning aids built around the three sourced clean-scan anchors, not a measured per-condition study, and they are flagged as such inside the instrument; a quantified residual curve across scan conditions is future work, not a result claimed here. The synthetic corpus brackets the real archive on the axes we identified and measured, but it cannot bracket variation we did not anticipate, so the transfer guarantee is conditional on the archive's variation staying inside the generator's coverage. The evaluation is against the native LAS of one public archive; transfer to archives with materially different scan vintages, track conventions or curve types would have to be revalidated, and the reality-gap residual is precisely the instrument for doing that revalidation rather than a promise that it is unnecessary.

References

1. Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., Abbeel, P. (2017). *Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World*. IROS. https://arxiv.org/abs/1703.06907

2. Tremblay, J., Prakash, A., Acuna, D., Brophy, M., Jampani, V., Anil, C., To, T., Cameracci, E., Boochoon, S., Birchfield, S. (2018). *Training Deep Networks with Synthetic Data: Bridging the Reality Gap by Domain Randomization*. CVPR Workshops. https://arxiv.org/abs/1804.06516

3. Ronneberger, O., Fischer, P., Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. MICCAI. https://arxiv.org/abs/1505.04597

4. Ganin, Y., Lempitsky, V. (2015). *Unsupervised Domain Adaptation by Backpropagation*. ICML. https://arxiv.org/abs/1409.7495

5 Jaderberg, M., Simonyan, K., Vedaldi, A., Zisserman, A. (2014). *Synthetic Data and Artificial Neural Networks for Natural Scene Text Recognition*. NeurIPS Deep Learning Workshop. https://arxiv.org/abs/1406.2227

6 Salehi, S. S. M., Erdogmus, D., Gholipour, A. (2017). *Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks*. MLMI. https://link.springer.com/chapter/10.1007/978-3-319-67389-9_44

7 Koroteev, D., Tekic, Z. (2021). *Artificial intelligence in oil and gas upstream: Trends, challenges, and scenarios for the future*. Energy and AI. https://www.sciencedirect.com/science/article/pii/S2666546820300033

Get the full whitepaper

This page is the long-form summary. The complete whitepaper adds the per-axis archive statistics behind each randomisation bound, the full per-curve and per-example R-squared, MAE and MSE tables under both the synthetic-validation and real-scan readings, the depth-resampling and LAS-alignment details of the evaluation protocol, and the worked reconciliation between synthetic validation and real-scan performance on the same held-out conditions.

EARTHSCAN

Sim-to-Real Transfer for Document-Style Geoscience Images

Published November 2022. Generated 2026-07-22.
https://www.earthscan.io/whitepapers/sim-to-real-transfer-document-style-geoscience-images