

淡江大學電機工程學系

指導教授：江正雄博士

神經微影模擬之成像品質分析與可靠度
提升

**Image Quality Analysis and Reliability
Enhancement in Neural Lithography
Simulation**

學生：詹盛惟
潘信瑞
沈群博

中華民國 115 年 10 月

摘要

神經微影模擬器可降低傳統物理模擬的計算成本，但現有方法多以 IoU、MSE 等全域像素指標衡量成像品質。對微影圖形而言，局部橋接或斷裂可能僅影響少量像素，卻會改變圖形連通性，因此高全域相似度未必代表模型能可靠預測近臨界缺陷。本研究以 LithoBench 為核心，探討神經微影模擬器在拓樸缺陷預測上的可靠度，並建立由邊緣帶 IoU、邊緣位移誤差至橋接與斷裂事件配對的三層評估流程。為建立具有鑑別力的測試條件，本研究對光罩施加形態學擾動，並透過物理模型重新模擬印出結果，以產生近臨界製程壓力案例。實驗首先分析全域成像指標是否足以反映缺陷可靠度，進一步比較原始資料、等量重複抽樣與物理擾動增強，以確認改善來源是否來自近臨界資料分佈。結果顯示，在維持 DOINN 架構與訓練設定的情況下，物理擾動增強使橋接 recall 由重複訓練基準的 0.520 提升至 0.8634，說明缺陷分佈覆蓋對模型可靠度具有明顯影響。在此基礎上，本研究提出拓樸感知模型 TopoNet，加入多尺度 Haar 小波細節、結構先驗、拓樸細化與非對稱拓樸損失，僅增加約 2.8% 的模型參數。TopoNet 的橋接 recall 進一步提升至 0.9457；在相同 571 個橋接參考事件上，相較物理擾動增強後的 DOINN，三次訓練平均漏檢數由 78 個降至 31 個，F1 由 0.9192 提升至 0.9454，但平均誤報亦由 8.67 個增加至 31.33 個。對 383 個斷裂事件，TopoNet 的 recall 為 0.8747。另一方面，在 194 個未擾動 cell 上，其 IoU 由原模型的 0.9216 降至 0.8207，具有拓樸誤報的切片由 1 個增加至 14 個，顯示缺陷敏感度提升與正常成像品質之間仍存在取捨。此外，本研究將方法延伸至 Ibex RISC-V 實體版圖，結合 DEF/LEF 的連線資訊，將部分幾何橋接事件進一步映射為跨網路電性短路風險。整體結果顯示，神經微影模擬器的可靠度不應僅由全域像素相似度判斷，近臨界資料覆蓋與拓樸感知設計皆能改善缺陷預測，但同時需量化其對正常成像與誤報所造成的代價。本研究建立一套可同時分析漏檢、誤報與電性風險的評估流程，作為後續可靠度導向神經微影模型設計的基礎。

關鍵詞：神經微影模擬、拓樸可靠度、橋接缺陷、製程壓力、TopoNet、電性短路

Abstract

Neural lithography simulators are commonly evaluated using global pixel-level metrics such as IoU and MSE. However, local bridge and break defects may affect only a small number of pixels while still changing pattern connectivity. This study investigates the defect-level reliability of neural lithography simulation using LithoBench and establishes a three-level evaluation framework based on edge-band fidelity, edge placement error, and matched topology events. Morphologically perturbed masks are re-simulated with a physics reference to generate near-critical process-stress cases. To examine whether reliability can be improved through data coverage and topology-aware modeling, the study compares repeated training, physics-based perturbation augmentation, and a proposed topology-aware model, TopoNet. Perturbation augmentation increases DOINN bridge recall from 0.520 to 0.8634, while TopoNet further improves it to 0.9457 with only a 2.8% increase in parameters. On 571 bridge events, the mean number of missed events decreases from 78 to 31, and F1 improves from 0.9192 to 0.9454, although false alarms also increase. TopoNet achieves a recall of 0.8747 on 383 break events. On 194 unperturbed samples, however, IoU decreases from 0.9216 to 0.8207, revealing a trade-off between defect sensitivity and nominal image quality. The method is further extended to an Ibex RISC-V layout, where DEF/LEF connectivity information is used to relate geometric bridge events to potential cross-net electrical shorts. Overall, the results show that neural lithography reliability should be evaluated beyond global image similarity, with both defect sensitivity and false-alarm cost considered.

Keywords: neural lithography modeling, topology reliability, bridge defects, process stress, TopoNet, electrical shorts

目錄

摘要.....	i
Abstract.....	ii
第一章 緒論.....	1
1.1 研究背景.....	1
1.2 問題陳述.....	1
1.3 研究目標與貢獻.....	2
第二章 相關研究與技術背景.....	3
2.1 微影成像與 OPC 驗證流程.....	3
2.2 逆微影技術與 SRAF.....	3
2.3 神經微影模擬器.....	4
2.4 傅立葉神經算子.....	4
2.5 小波分解與多解析度分析.....	4
2.6 版圖驗證與 hotspot 偵測.....	5
2.7 文獻缺口與本研究定位.....	5
第三章 研究方法.....	6
3.1 整體架構.....	6
3.2 缺陷層級評估框架.....	7
3.3 事件計數方式.....	8
3.4 物理重模擬的形態學壓力集.....	8
3.5 形態學擾動與劑量掃描的覆蓋率對齊.....	9
3.6 物理擾動增強.....	10
3.7 拓樸感知架構 TopoNet.....	10
3.8 真實電路的驗證鏈.....	12
3.9 電性短路的判定.....	15
3.10 功能失效的推導.....	16
第四章 實驗結果與分析.....	17
4.1 實驗環境與訓練設定.....	17
4.2 資料集.....	17
4.3 全域指標無法預測缺陷偵測能力.....	18

4.4 盲區集中於臨界製程壓力區	20
4.5 低擾動餘裕位置的漏檢風險	21
4.6 擾動餘裕的系統性高估	22
4.7 診斷：模型端修改的改善有限	22
4.8 診斷證據：近臨界訓練案例不足	24
4.9 解法一：物理擾動增強	24
4.10 排除混淆變因	25
4.11 解法二：TopoNet	27
4.12 消融實驗	30
4.13 真實電路的驗證設定	31
4.14 通過路由規則檢查版圖的劑量方向餘裕	33
4.15 失效的幾何選擇性	34
4.16 單一幾何間距門檻的篩選限制	35
4.17 修復方法的免重訓遷移	36
4.18 Ibex 的 precision、recall 與閾值取捨	37
4.19 跨層與跨設計的一致性	38
4.20 電性短路的失效型態	39
4.21 形態學製程壓力與直接劑量掃描	40
4.22 印出閾值變動下的模型表現	41
4.23 形態學擾動與劑量掃描的涵蓋關係	41
4.24 研究限制	43
第五章 結論與未來發展	45
5.1 結論	45
5.2 未來發展	46
參考文獻	47

圖目錄

圖 1. 本研究兩階段流程、MetalSet 同一樣本之 mask-to-print 流程與 morphology stress 生成示意	6
圖 2. TopoNet 概念架構與訓練期非對稱損失	11
圖 3. RTL 至短路風險分析流程、OpenROAD 切片與 NeuralILT 跨域範例	13
圖 4. 代表性 Ibex 切片中，原始梯度 ILT 與 topology-constrained ILT 的光罩／印出結果對照	15
圖 5. LEF pin 與 raster 版圖對齊，以及幾何橋接轉換為 electrical-short 的判定概念	16
圖 6. 同一 StressMetal 樣本的多模型輸出、全域 IoU 與 bridge recall 關係，以及 edge-band 定義	19
圖 7. 光罩形態學擾動下的事件配對 recall 與原始事件計數	20
圖 8. 依首次失效擾動餘裕分級的 hotspot 偵測率	21
圖 9. 模型端修改的各自對照與尺度遞迴架構的 mIoU	23
圖 10. 抽樣元件間隙與模型 EPE 標準差的尺度比較	24
圖 11. CtrlSet 與 MixSet 對照：改善主要來自資料分佈，而非單純增加樣本量	26
圖 12. 8 倍平均池化與 Haar 小波對細縫高頻資訊的保留差異	27
圖 13. TopoNet 各機制的單一 seed 消融結果	31
圖 14. OpenROAD 產生之 Ibex RISC-V routed layout 全晶片視圖	32
圖 15. Ibex metal2 代表性 electrical-short 案例：design／net、物理真值與三種神經模擬器輸出	34
圖 16. 跨層與跨設計的 electrical-short recall 比較	38
圖 17. Morphology stress 與 direct-dose stress 的 shared failure 與 mismatch 代表案例	42

表目錄

表 1. 三層評估架構.....	7
表 2. 訓練集印出結果的元件間隙（500 個樣本、1983 個元件配對）	9
表 3. 形態學擾動與直接劑量的印出覆蓋率對齊	10
表 4. TopoNet 的四個機制	11
表 5. 解碼器各層的輸入通道數	12
表 6. 拓樸約束對光罩品質的影響（1200 個切片）	14
表 7. 功能失效型態的分類	16
表 8. 實驗環境與訓練設定	17
表 9. 資料池與評估資料規模	18
表 10. 四個模擬器的缺陷層級表現	18
表 11. 全域／邊緣帶 IoU 與缺陷 recall 的 Spearman 等級相關（n = 39 模型變 體）	19
表 12. 依形態學擾動分組的事件計數與配對 recall	20
表 13. 依風險等級的 hotspot 偵測率	21
表 14. 擾動餘裕的高估量	22
表 15. 四組模型端修正方法的結果	23
表 16. 物理擾動增強的效果	25
表 17. 資料池規模與每 epoch 梯度步數對照	25
表 18. 擾動泛化測試	26
表 19. 隨機種子變異	27
表 20. 架構比較（皆於 MixSet 訓練）	28
表 21. 橋接事件配對結果（每列 571 個參考事件）	28
表 22. 斷裂事件配對結果（每列 383 個參考事件）	29
表 23. 194 個未擾動 cell 的成像品質與誤報	29
表 24. 消融實驗結果	30
表 25. 受測電路	31
表 26. Ibex 版圖的劑量方向製程餘裕分佈（740 處）	33
表 27. 劑量方向製程餘裕最小的位置	33
表 28. 依設計間距分組的失效率	35

表 29. 單一幾何門檻的涵蓋率與誤報占比（86 個參考事件）	36
表 30. LithoBench 與 Ibex 原始 ILT 主測試的 recall	36
表 31. precision 與 recall（1200 切片）	37
表 32. 判定閾值掃描（200 切片）	37
表 33. 跨層與跨設計的 recall（標示光罩版本）	38
表 34. 可有效分類之功能失效型態（1135 事件）	39
表 35. 依功能失效型態分類的偵測率	40
表 36. 形態學製程壓力與直接劑量掃描的比較	40
表 37. 物理印出閾值變動下的 recall（500 個切片）	41
表 38. 兩種變異方式的涵蓋關係（300 個切片）	42

第一章 緒論

1.1 研究背景

積體電路製造透過微影製程將光罩圖形轉印至晶圓。在深紫外 (DUV) 微影中，193 奈米 ArF 為常見曝光波長之一；當圖形尺寸接近或低於光學系統可解析尺度時，繞射與製程效應會造成圓角、線端縮短及線寬偏移 [14]。計算微影因此採用光學鄰近效應修正 (Optical Proximity Correction, OPC) 與逆微影技術 (Inverse Lithography Technology, ILT)，反覆調整光罩，使成像結果更接近設計目標 [9][10][14]。每一次光罩更新都需要成像評估，這也是神經微影模擬器可協助降低計算需求的環節。設計規則檢查 (Design Rule Check, DRC) 依規則集檢查 drawn layout，實際規則可包含幾何與圖形上下文；微影驗證則利用成像模型評估指定條件下的印出輪廓。通過某一規則集不直接代表已驗證所有曝光、焦距與光阻條件下的成像結果。本文使用公開物理模型作為共同參照，不將其直接對應至特定代工廠製程節點，並以首次出現拓樸失效的擾動幅度描述相對餘裕。

1.2 問題陳述

微影成像預測需要同時回答「輪廓是否相似」與「關鍵連通性是否相同」。神經微影模型常以全域 IoU、MSE 等指標呈現成像品質 [1][2][3]，但橋接可能只改變狹窄間隙，斷裂也可能只影響少量細頸像素；這些位置的錯誤對全域分數影響有限，卻會改變拓樸事件判定。以本研究的 MetalSet 正常測試為例，DOINN 的 IoU 可達 0.949，因此仍需另外檢驗它對接近失效邊界的光罩能否忠實預測。本研究的主題是神經微影模擬器相對物理參考的缺陷預測可靠度。物理重模擬用於建立訓練與測試標記；神經模型則由光罩直接預測成像。本文以受控擾動檢查漏判、提出改善並量測代價，實際 ILT 迭代加速及晶圓良率列為後續應用問題。Ibex 延伸分析則用於說明幾何缺陷如何對應到電路連線風險。

1.3 研究問題

本文以三個可由實驗回答的研究問題組織核心分析，分別對應第 4.2、第 4.3 與第 4.4 節。

Q1：全域成像相似度是否足以反映缺陷可靠度？以跨模型指標比較、擾動幅度分組及低餘裕事件評估，檢查全域 IoU 與橋接、斷裂等局部連通性判斷的關

係。

Q2：橋接漏判改善是否來自缺陷分佈的補充？在維持 DOINN 架構與訓練設定的條件下，比較原始資料、等量重複抽樣與物理擾動增強，並以擾動涵蓋測試檢查對訓練分佈的依賴。

Q3：拓樸感知機制可帶來多少改善及何種代價？在共同 MixSet 上比較增強 DOINN、WRD-Net 與 TopoNet，量測事件 precision/recall/F1、正常樣本 IoU 與誤報，並以消融觀察各機制的作用。

1.4 研究貢獻與章節安排

第一，建立由像素、邊緣到事件配對的評估流程。形態學光罩由物理模型重新產生參考輸出，使正常成像品質與近臨界缺陷可靠度能分開量測，並將橋接、斷裂及誤報定位到相應的事件單位。第二，以資料量控制與擾動涵蓋實驗支持近臨界分佈的重要性。物理擾動增強使橋接 recall 由既有重複訓練基準 0.520 提升至 0.8634，提供可在基準架構上實施的改善途徑。第三，提出 TopoNet 並量化其改善與取捨。相對增強 DOINN，三次訓練的平均橋接 recall 增加 8.23 個百分點、F1 增加 2.62 個百分點；同時揭露正常樣本的成像與誤報代價，使「可靠度提升」具有明確的條件與量測依據。延伸驗證使用 Ibex、不同金屬層及 gcd，檢查免重訓的改善方向，並透過 DEF/LEF 將部分幾何事件映射至跨網路短路風險。此部分作為連接電路設計的實例，與 LithoBench 核心比較分別報告。第二章說明相關研究與技術定位；第三章定義資料、事件與模型方法；第四章先回答三個研究問題，再呈現 Ibex 延伸、代理有效範圍及綜合討論；第五章收斂主要結論與後續研究。

第二章 相關研究與技術背景

2.1 微影成像與 OPC 驗證流程

在部分同調照明下，光罩圖形到晶圓空中像的映射可由 Hopkins 成像理論描述。實務數值計算常將傳輸交叉係數（Transmission Cross Coefficient, TCC）進行特徵分解，保留主要模態形成 Sum of Coherent Systems（SOCS）近似，使空中像可表示為多組光學核與光罩卷積後強度的加權和，再經光阻模型得到印出輪廓 [14][32][33][34][35]。此類模型可直接描述光學核、焦距與劑量等條件，因此適合作為神經 surrogate 的物理參照。在驗證流程中，DRC、MRC 與微影驗證分別處理設計版圖、光罩可製造性與製程後印出結果。drawn geometry 通過幾何規則檢查，不代表其 process printing 具有足夠餘裕。本研究第 4.5.4 節因此以單一 spacing threshold 作為簡化幾何 baseline，用於評估此類門檻能否重現物理模擬所辨識的失效，而不將其視為完整 foundry DRC 的替代方案。

2.2 逆微影技術與 SRAF

Model-based OPC 透過反覆模擬與邊緣修正降低目標圖形與印出圖形之差異；ILT 則將光罩像素或輪廓視為最佳化變數，將微影成像表述為逆問題，並以像素誤差、EPE、製程變異帶（PV band）或光罩複雜度等項目建立目標函數 [9][10]。近年的公開 ILT 研究涵蓋 GPU level-set、multi-level simulation、深度神經 level-set 與開源 ILT framework 等方向 [16][17][18][23]；全晶片尺度仍高度依賴平行運算與分塊處理 [22]。相較於傳統局部邊緣修正，ILT 具有較大的解空間，也更需要處理光罩可製造性與拓樸正確性。SRAF 利用不希望獨立印出的輔助圖形調整鄰近主圖形的成像場，使主圖形在不同製程條件下更穩定 [32][34]。GAN-OPC、Neural-ILT、AdaOPC 與 physics-informed ILT 等方法則以神經網路預測、初始化或輔助最佳化光罩，以降低逐圖反覆迭代的成本 [5][19][21][24][25][26]；LithoBench 中的 NeuralILT 實作亦採 encoder-decoder 與 skip connection 設計，可追溯至 U-Net 類架構 [40]。本研究第二階段使用可微分物理模型產生梯度式 ILT 光罩，其用途為建立可控制的實驗輸入，未將其定位為商用 OPC/ILT/MRC flow 的替代方案。

2.3 神經微影模擬器

LithoBench 提供統一的資料、物理真值與訓練／評估介面，用於比較深度學習式微影模擬與光罩最佳化方法 [1]。本研究採用其官方框架中的 DOINN、DAMOLitho、CFNOLitho 與 LithoGAN 相關實作作為代表性模型。LithoGAN 以條件式生成對抗網路進行 mask-to-resist 映射 [3]；DOINN 結合全域頻域感知與局部卷積感知，利用傅立葉算子捕捉較長距離的光學關聯 [2]；CFNO 則以卷積式 Fourier operator 引入 lithography-aware inductive bias [20]。後續 LithoSim 與 OptiCo 等研究亦持續朝更大規模資料、process variation 與顯式物理先驗整合發展 [27][28]。上述方法的架構雖不同，常見評估仍集中於 MSE、pixel accuracy、IoU、EPE 或整體輪廓誤差 [1][20][25][28]。這些指標可衡量全局保真度，但未直接量測預測輪廓在正確位置發生橋接、斷裂或圖形消失的能力。因此，本研究著重於補足 defect-level reliability 的評估層次。

2.4 傅立葉神經算子

Fourier Neural Operator (FNO) 將積分算子的 kernel 參數化於傅立葉域，利用頻域乘法在單層中取得全域 receptive field，因此適合描述具有長距離交互作用的物理場映射 [6]。DOINN 將此概念引入微影模擬，使全域頻域分支處理較低頻、較長距離的光學資訊，再由局部分支保留細節 [2]。需要指出的是，本研究所觀察到的細縫資訊衰減來自所採 DOINN 實作的具體資料路徑，不能概括為 FNO 的必然限制：其全域分支在進入頻域模組前先進行 8 倍平均池化，當 gap 尺度接近或小於此下採樣尺度時，局部高頻差異會顯著衰減。第 3.7 節與第 4.4 節據此進行 TopoNet 的結構診斷與設計。

2.5 小波分解與多解析度分析

多解析度分析將影像分解為不同尺度與頻帶，使低頻結構與高頻邊緣可分別處理。二維 Haar 小波可將每個尺度分解為 LL、LH、HL 與 HH 四個子帶；其中 LL 對應局部平均的低頻資訊，而其餘三個 detail band 保留水平、垂直與對角方向的局部變化 [12]。這個性質與本研究的問題高度吻合：平均池化能保留大面積主圖形，卻可能削弱狹窄 gap；Haar 分解則可在保留低頻路徑的同時，額外輸出 detail band。TopoNet 因此保留原本的全域分支，並將多尺度細節作為輔助通道注入 decoder，以補回下採樣後弱化的高頻拓撲線索。另一方面，影像分割

領域已有多種 topology-aware objective 指出：對細線、網路或連通結構而言，單純 pixel-wise loss 可能低估斷裂與錯接的影響；相關方法包括高階拓樸特徵損失、Betti-number/persistent-homology 約束與 skeleton-based cIDice [36][37][38][39]。這些方法並非直接針對微影而設計，但共同支持「連通性錯誤應被獨立於像素誤差量測」的觀點，亦構成本研究 topology-level evaluation 的方法論背景。

2.6 版圖驗證與 hotspot 偵測

微影 hotspot 偵測方法大致可分為 pattern matching 與 machine-learning/deep-learning 類型；後者通常直接以設計版圖或其特徵為輸入，預測 pattern 的高風險程度 [13]。本研究著重於另一層次的問題：先以物理微影模型產生 reference printed pattern，再將神經模擬器預測與 reference 進行事件級比對；在真實電路階段，進一步利用 DEF/LEF 的 net 資訊判斷幾何橋接是否構成 electrical short。因此，本文評估的是 surrogate simulator 的可靠度，與直接訓練 layout-hotspot classifier 的任務設定不同。

2.7 文獻缺口與本研究定位

本文將研究焦點放在公開神經微影基準的缺陷預測可靠度。DOINN 與 LithoBench 提供成像建模及統一比較的基礎 [1][2]；影像拓樸損失與微影 Graph Loss 已指出像素相似度不能完整描述連通性 [15][36][37][38][39]。本研究在此基礎上，進一步量測近臨界光罩的事件漏判，透過資料量對照檢查改善來源，並同時呈現事件品質與正常輸出的取舍。本文的貢獻以這組問題、受控比較及評估結果為界，不將使用拓樸指標本身視為獨有概念。

第三章 研究方法

3.1 整體架構

(a) 本研究兩階段流程



(b) 同一 MetalSet 樣本：上排為目標、光罩、物理參考；下排為原版、增強、TopoNet。



(c) 形態學壓力範例：上排為擾動光罩，下排為重新計算的物理印出結果。



圖 1. 本研究兩階段流程、MetalSet 同一樣本之 mask-to-print 流程與 morphology stress 生成示意

本研究的整體流程如圖 1 所示。第一階段以 LithoBench 為主要實驗平台，建立成像與事件級評估、診斷漏判並提出資料與架構改善；第二階段保留 Ibex 等版圖的延伸驗證，將幾何連通性變化映射至跨網路短路風險。本研究將兩個階

段分別用於幾何／拓樸保真度與電性後果評估。第一階段使用 LithoBench MetalSet；此資料集由版圖切片構成，僅含幾何形狀，不含電路連線資訊，因此用於量測神經模擬器預測與物理真值的一致性。第二階段使用 OpenROAD 產生的處理器版圖，並利用設計交換格式（DEF）中的連線名稱評估預測誤差是否進一步形成 electrical short。圖 1(b)以同一 MetalSet 樣本呈現 mask、物理真值與神經模擬輸出的視覺差異；圖 1(c)示意受控 morphology stress 的生成方式。bridge、break 與 disappearance 三類拓樸事件於第 3.2 節進一步定義。

3.2 缺陷層級評估框架

本研究建立的評估框架分為三個層級，如表 1 所示。

表 1. 三層評估架構

層級	量測內容	用途
L1 邊緣帶	IoU 限制在輪廓外擴 8 像素的窄帶內	與全域指標對照
L2 EPE	沿輪廓量測印出邊緣相對目標的位移	ILT 品質與模擬器保真度
L3 拓樸事件	連通元件分析：橋接、斷裂、消失、SRAF 誤印	缺陷偵測能力

L3 的橋接判定為本框架的關鍵設計。初版以「預測事件數／真值事件數」作為比值指標，但此方法未考慮事件位置，可能使事件數相近但位置錯誤的預測取得偏高分數。修正版改採配對比對：以未擾動印出結果作為拓樸參照，找出落入同一連通元件的元件對，分別於真值與預測中建立集合，再由集合運算取得 hit、miss 與 false，如公式(1)所示。

$$\text{hit} = |\text{pairs}(Z_o, \text{gt}) \cap \text{pairs}(Z_o, \text{pred})| \quad (1)$$

令真值事件集合為 G 、預測事件集合為 P ，則 $TP = |G \cap P|$ 、 $FN = |G - P|$ 、 $FP = |P - G|$ 。本文以 $\text{precision} = TP/(TP+FP)$ 、 $\text{recall} = TP/(TP+FN)$ 、 $F1 = 2TP/(2TP+FP+FN)$ 量測事件配對品質。無預測事件時 precision 不定義；本文亦未建立可枚舉的無缺陷事件集合 TN ，因此不將 $FP/(TP+FP)$ 稱為 false positive rate (FPR)，而稱為誤報占比或 $1-\text{precision}$ 。

3.3 事件計數方式

在真實電路版圖中，元件配對計數會系統性放大同一融合區所造成的事件

數。若一個連通元件涵蓋 N 條連線，配對計數將產生 $C(N,2)$ 個事件，而這些事件實際對應同一處金屬融合區，彼此高度相關。在較大擾動下，Ibex 版圖的配對計數為 5781 個事件；改以融合區計數（一個連通元件涵蓋兩條以上不同連線記為一處）後僅有 1182 處，其中 4599 個屬重複配對。配對計數亦使偵測率趨勢由修正後的遞增轉為遞減，與 benchmark 結果不一致。基於此觀察，本研究依資料來源採用兩種事件定義：MetalSet 等匿名幾何 benchmark 以元件配對進行 bridge event matching；具有 net 標籤的真實電路則以融合區（fusion region）定義 electrical-short site，以避免同一多-net 連通區被 $C(N,2)$ 重複計數。兩種定義分別服務於幾何保真度與實體失效位置的量測。斷裂評估另區分原始 brk_gt 計數、含斷裂的切片數與事件配對結果。整體斷裂 precision/recall/F1 使用同一次事件配對輸出的 TP、FP、FN；原始計數不能直接拿來作此 recall 的分母。本文保留這些層次的名稱與數量，避免以不同輸出相除。

3.4 物理重模擬的形態學壓力集

LithoBench 的真值由 ILT 最佳化光罩產生，因此原始資料集中的拓樸事件極少。為量化訓練分佈與臨界缺陷之間的差距，本研究直接量測其印出結果，統計如表 2 所示。

表 2. 訓練集印出結果的元件間隙（500 個樣本、1983 個元件配對）

統計量	間隙（像素）
最小	35
第 1 百分位	53
第 5 百分位	65
中位數	104
小於 5 像素的配對	0
小於 10 像素的配對	0

模型的邊緣位移誤差為 -0.694 ± 1.102 像素。相較之下，500 個 MetalSet 印出樣本中的 1983 個元件配對最小間隙為 35 像素；將元件面積門檻 min_area 分別設為 20、100、200、500 像素時，配對數、最小值、第 1 百分位 53 像素與中位數 104 像素皆一致。此數字是抽樣結果，EPE 標準差亦不是最大誤差上界，因此不能據此推論橋接在數學上不可能發生。為提高評估鑑別力，本研究對 ILT 光罩施加膨脹或收縮，以物理模擬器重新計算輸出，並保留出現拓樸事件的

案例。未擾動光罩重新模擬後，與資料集內附印出結果的 IoU 及像素準確率皆為 1.0000，確認本管線可重現同一 nominal reference。

3.5 形態學擾動與劑量掃描的覆蓋率對齊

形態學擾動以像素為單位，與物理曝光劑量屬不同變異方式。本研究分別量測不同擾動量與 direct-dose sweep 所造成的印出覆蓋率，並以覆蓋率相近的操作點建立相對校正關係。此校正僅用於估計壓力尺度，不假設兩種變異在空間位置上完全等價。校正結果與圖案類型相關：MetalSet 塊狀金屬的 +2 px 約對齊 +6.6% dose；Ibex 細線圖形的 +1 px 約對齊 +7.0%，+2 px 約對齊 +16.5%。因此本文依版圖類型分別標示低幅度主評估與較強壓力操作點，如表 3 所示。

表 3. 形態學擾動與直接劑量的印出覆蓋率對齊

資料集	擾動量	覆蓋率對齊之劑量偏移	本研究用途
MetalSet	+2 px	+6.6%	低幅度主評估
MetalSet	+4 px	+15%	較強壓力
Ibex metal2	+1 px	+7.0%	低幅度主評估
Ibex metal2	+2 px	+16.5%	較強壓力

3.6 物理擾動增強

基於近臨界缺陷樣本不足的診斷，本研究以與壓力集相同的方式產生含缺陷訓練樣本，並混入原始訓練集。此資料增強策略維持網路結構、參數量與訓練超參數不變。為避免資料來源重疊，擾動樣本取自訓練集所屬的前 90% 版圖，測試用壓力集則取自後 10%，混合後再進行隨機打散，並確認原測試集未進入訓練側。

3.7 拓樸感知架構 TopoNet

本研究進一步診斷基準模型 DOINN 的結構瓶頸。其全域分支的實作如公式 (2) 所示：

$$\text{br}_0 = \text{RFNO}(\text{avg_pool2d}(x, 8, 8)) \quad (2)$$

所採 DOINN 實作的全域頻域分支約有 131K 參數，約佔全模型的 62%，且在 RFNO 前先進行 8 倍平均池化。壓力集中的近臨界細縫可小至數個像素，因此此路徑可能削弱局部高頻差異。TopoNet 保留 DOINN 主路徑，另加入多尺

度細節、結構先驗、拓樸細化與非對稱損失；圖 2 呈現概念資料路徑，表 4 說明四個機制的作用。

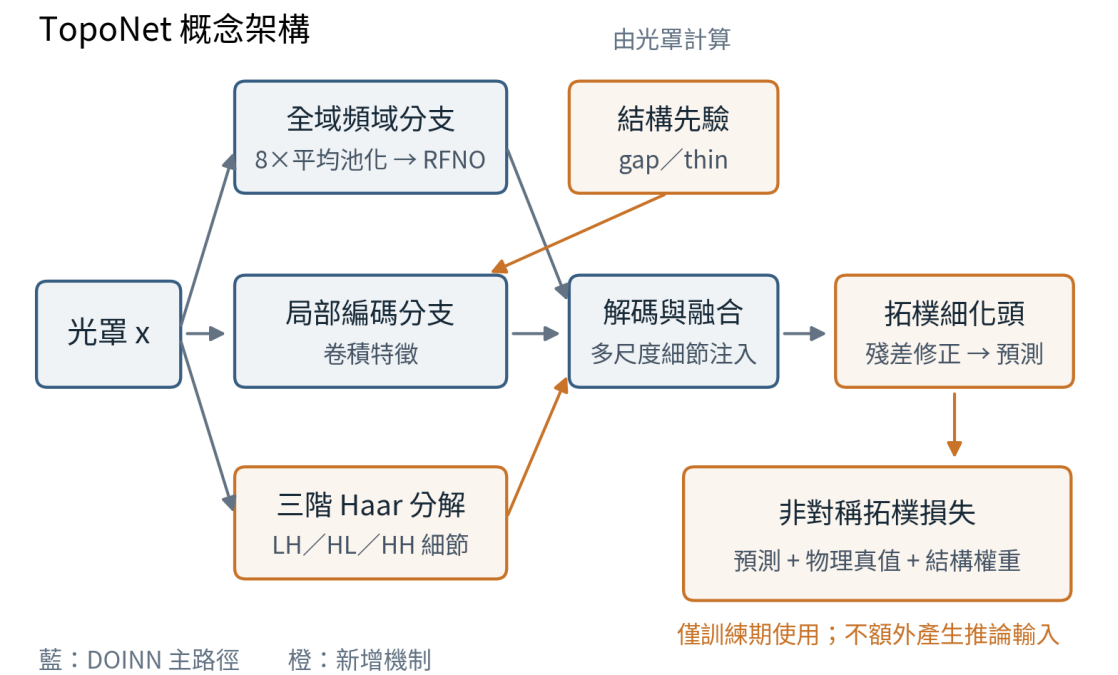


圖 2. TopoNet 概念架構與訓練期非對稱損失

表 4. TopoNet 的四個機制

機制	作用
三階 Haar 小波分解	保留 LH/HL/HH 細節，注入多尺度解碼器
結構先驗注入	由光罩計算 gap/thin 圖，提供局部結構資訊
拓樸細化頭	以殘差修正預測；不保證只影響連通性
非對稱拓樸損失	對 gap 欠預測及 thin 過預測加重懲罰

合計參數量為 217.6K，相對基準模型的 211.6K 增加 2.8%。各解碼層的輸入通道數變化如表 5 所示，其中每層增加的 3 個通道即為小波的細節子帶。

表 5. 解碼器各層的輸入通道數

解碼層	輸入組成	DOINN	TopoNet
deconv0	RFNO 64 + 編碼器 64 + 細節子帶 3	128	131
deconv1	前層 32 + 編碼器 32 + 細節子帶 3	64	67
deconv2	前層 16 + 編碼器 16 + 細節子帶 3	32	35

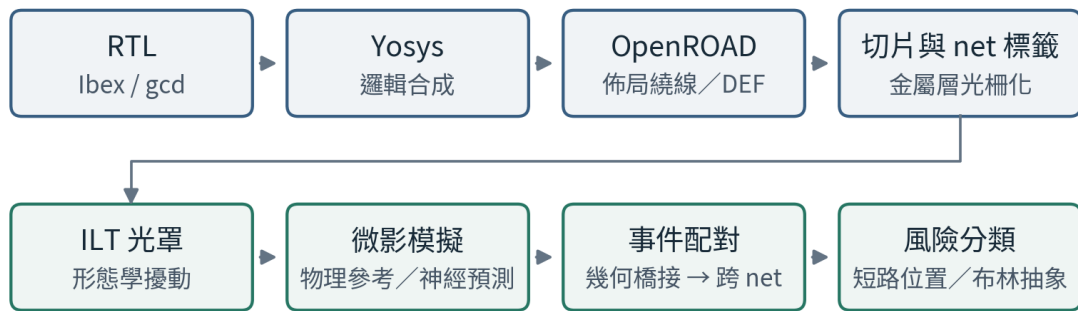
非對稱損失針對缺陷漏判的誤差方向設計：在真值已橋接的位置，欠預測可能保留錯誤間隙；在真值已斷裂的位置，過預測可能錯誤補回細頸。公式(3)以 $\text{err} = \text{pred} - \text{gt}$ 表示預測誤差，透過 gap 與 thin 權重提高相應區域的訓練懲罰。此權重不是對所有橋接／斷裂情況的保證。

$$w = 1 + \lambda_g \cdot \text{gap} \cdot 1[\text{err} < 0] + \lambda_t \cdot \text{thin} \cdot 1[\text{err} > 0] \quad (3)$$

3.8 真實電路的驗證鏈

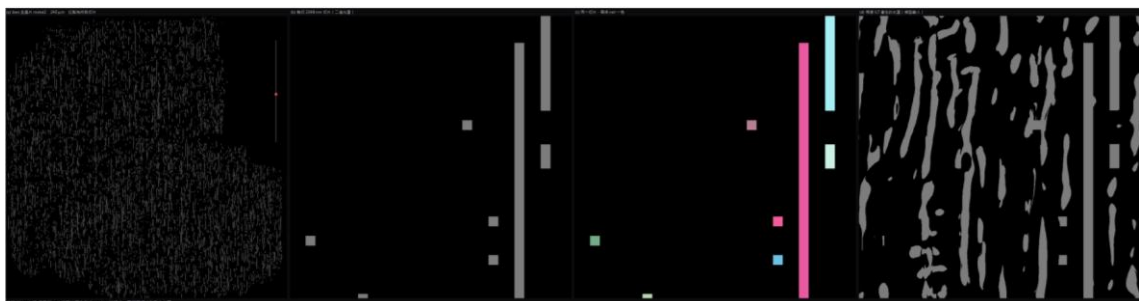
第二階段的流程如圖 3 所示。以 Yosys 進行邏輯合成 [31]、OpenROAD 進行佈局與繞線 [7]，其中 detailed routing 由 TritonRoute 系列方法支援，並透過 pin-access 與 DRC-aware routing 處理實體連線 [29][30]，產生 Ibex RISC-V 處理器的完整版圖。設計交換格式 (DEF) 中記錄每條連線的名稱與所有金屬片段的座標，本研究由此抽取指定金屬層並光柵化為模擬器可接受的格式。圖 3(b)進一步示意由完整版圖裁切、net 標籤對齊到 rasterization 的資料轉換；圖 3(c)則呈現 NeuralILT 直接套用於 Ibex 的代表性輸出。

(a) 從 RTL 到連線與邏輯風險的分析流程



跨網路連通性與邏輯風險分析；完整電性／功能模擬列為後續工作

(b) OpenROAD 完整版圖 → 切片 → net 標籤 → 光柵化／光罩處理。



(c) NeuralILT 跨域範例：目標、預測光罩、物理印出、差異圖（紅：缺失；藍：額外）。



100 個切片的 target coverage：平均 85.7%、中位數 87.3%；此單例為 87.4%。

圖 3. RTL 至短路風險分析流程、OpenROAD 切片與 NeuralILT 跨域範例

光罩產生方面，本研究將 LithoBench 的 MetalSet NeuralILT 權重直接套用於 100 個 Ibex metal2 切片，並以相同物理模擬器重新計算印出結果。target coverage 平均為 85.7%、中位數為 87.3%，顯示 NeuralILT 在此設定下仍具有一定程度的跨資料域可印刷性。第二階段仍採用可微分物理模型進行梯度式 ILT，主要原因是其目標函數可直接加入 net-aware gap protection 與 topology constraint，便於建立可控制、可重現的拓樸安全光罩生成流程。target coverage 僅衡量目標區域被印出的比例，無法代表 topology correctness、額外印刷量或製程窗口品質。基礎目標函數如公式(4)所示。圖 3(c)中的單一案例僅供視覺化；85.7%/87.3% 的 coverage 結論來自 100 個切片的整體統計。

$$\min \text{MSE}(\text{printed}(\text{mask}), \text{target}) + \lambda \cdot \text{PV}_{\text{band}} \quad (4)$$

最佳化結果在主圖形之外出現 70 個獨立小區塊，依本研究光柵化尺度換算的面積中位數為 188 平方奈米（等效邊長約 14 奈米）。其中部分小區塊在 nominal 物理模擬中未形成獨立印出圖形，形態上與輔助成像用途相近 [32][34]，故本文稱為 SRAF-like feature；其實際光學貢獻與光罩可製造性未經完整驗證。另一方面，僅含 MSE 與 PV band 的目標函數未納入 net-level 拓樸資訊，初版光罩在 1200 個切片中有 8 處於標稱劑量下形成跨網路短路。本文將其歸類為實驗光罩生成的 nominal artifact，並於目標函數中加入公式(5)的間隙保護項。

$$L = \text{MSE}(\text{printed}, \text{target}) + \lambda_{\text{pv}} \cdot \text{PV}_{\text{band}} + \lambda_{\text{gap}} \cdot \text{mean}(\text{relu}(\text{printed}[\text{gap}] - \tau)^2) \quad (5)$$

gap 區域由連線標籤直接計算。在 1024×1024 的處理解析度下，不同 net 的相鄰圖形之間、 $\text{max_gap} = 140 \text{ px}$ 範圍內的背景像素均納入保護，依本文光柵化尺度約對應 280 奈米；same-net 間隙不納入保護。該區域佔全圖面積的中位數為 4.12%。加入約束後，印出 IoU 由 0.784 降至 0.660，顯示 nominal

topology protection 與逐像素保真度之間存在明確取舍。

加入該約束後的結果如表 6 所示。圖 4 進一步以代表性切片呈現原始 gradient ILT 與 topology-constrained ILT 在 mask 與 printed pattern 上的差異，對照可見 topology protection 主要抑制 nominal condition 下的跨網路橋接。

表 6. 拓樸約束對光罩品質的影響 (1200 個切片)

光罩	印出 IoU 中位	IoU 最低	標稱劑量下的短路
僅 MSE + PV band	0.784	0.586	8 處
加入間隙保護項	0.660	0.435	0 處

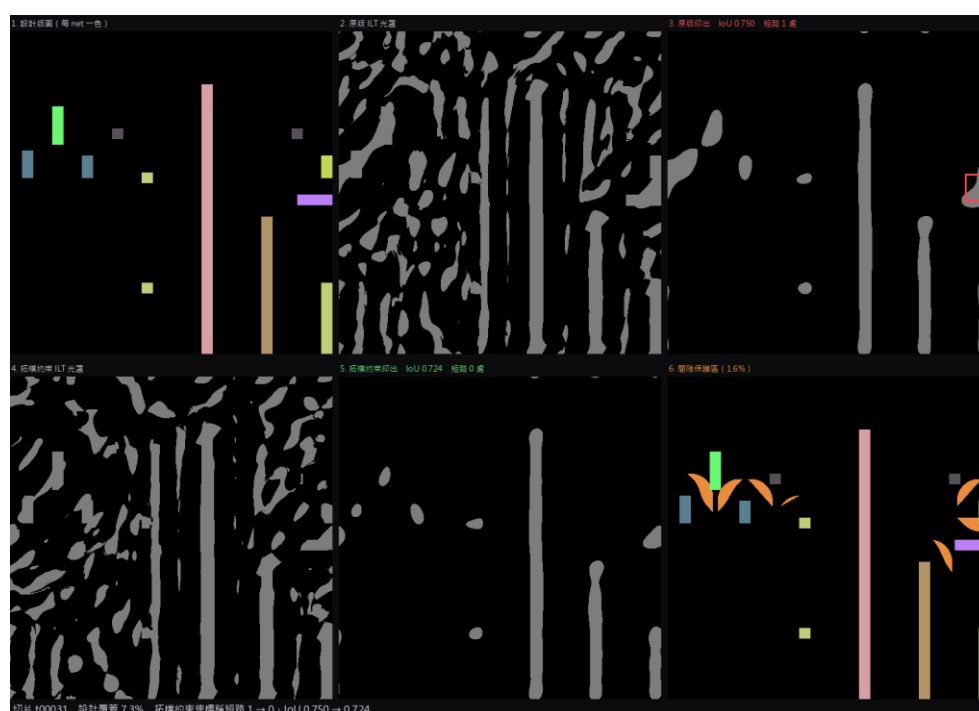


圖 4. 代表性 Ibex 切片中，原始梯度 ILT 與 topology-constrained ILT 的光罩／印出結果對照

加入拓樸約束後，標稱劑量下的電性短路由 8 處降至 0 處，同時印出 IoU 由 0.784 降至 0.660。此結果顯示像素相似度與拓樸正確性並非完全等價。topology-constrained 版本主要作為 nominal-topology control，用於確認後續模型比較未受 8 個 mask-generation artifact 主導。

3.9 電性短路的判定

幾何橋接需進一步結合 net 資訊才能判定是否構成電性短路。當兩塊相連金屬分屬不同電路連線 (net) 時記為 short；same-net merge 不列入 electrical-short 統計。圖 5(b)示意此判定原則。在 Ibex 版圖中，DEF 提供完整的連線名稱與座標，涵蓋率為 100%。標準元件則由 Nangate45 LEF 檔取得各 pin 在 metal1 層

的完整多邊形，並與版圖座標對應；40 個元件的實測座標關係均一致，其換算如公式(6)所示。圖 5(a)提供 LEF pin polygon 與光柵化版圖的實際對齊範例。

$$\text{影像座標} = \text{LEF 座標} \times 1000 + (0, dy) + (384, 384) \tag{6}$$

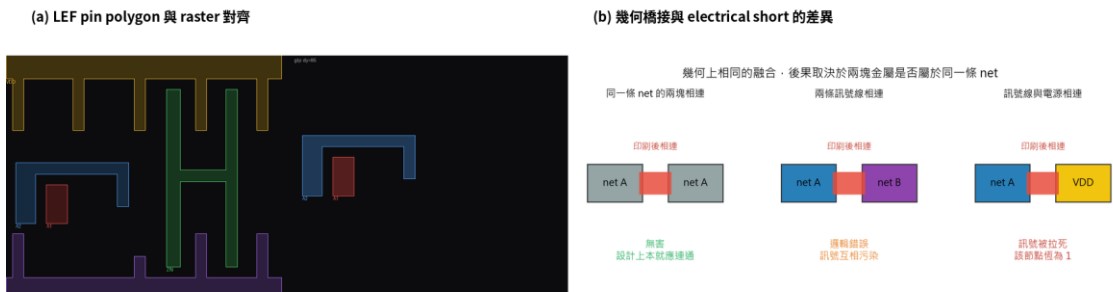


圖 5. LEF pin 與 raster 版圖對齊，以及幾何橋接轉換為 electrical-short 的判定概念

短路風險依所涉及的連線類型區分：電源與電源之間的橋接可能造成高電流或供電異常，屬最高風險；訊號與電源橋接可能形成 stuck-like 行為或驅動衝突；訊號與訊號橋接則可能改變組合或循序邏輯。實際電性後果仍取決於橋接電阻、驅動能力與電路狀態，本研究在此僅進行 net-level 與數位邏輯層級分類。

3.10 功能失效的推導

Nangate45 的標準元件資料提供可機器解析的布林功能。本研究將短路的兩個 pin 在數位邏輯層級抽象為共享同一邏輯變數，重新展開真值表後與原始邏輯比對，據此將事件分類為不同功能失效型態，如表 7 所示。此分析屬於邏輯層級分類，未包含橋接電阻、驅動強度與瞬態效應等類比 short-circuit 行為。

表 7. 功能失效型態的分類

失效型態	說明	範例
降維	多輸入閘的輸入短接，等效輸入數減少	AND3 的 A1-A2 短接
退化	二輸入閘退化為緩衝器	AND2 $\rightarrow$ ZN = A1
輸入釘死	輸入被拉至邏輯 0 或 1	A1 與 VSS 短路
輸出釘死	輸出恆為固定值	ZN 與 VDD 短路
輸出回授	輸出與輸入短接，形成組合迴路	ZN 與 A1 短路

第四章 實驗結果與分析

4.1 實驗設計與比較原則

本章依序回答三個研究問題：全域成像指標是否足以描述缺陷可靠度、資料分佈改善是否超過等量擴充，以及 TopoNet 帶來多少缺陷改善與正常成像代價。核心比較集中於 LithoBench；Ibex 與製程代理分析用於檢查延伸用途及適用範圍。

4.1.1 實驗環境與訓練設定

表 8. 實驗環境與訓練設定

項目	設定
GPU	NVIDIA GeForce RTX 3080 Ti (12 GB)
深度學習框架	PyTorch / LithoBench
影像解析度	1024 × 1024
Batch size	4
最佳化器	Adam
初始學習率與排程	1e-3，半程 StepLR 衰減 0.1 倍
損失函數	MSE(aerial, litho) + MSE(resist, printed)
訓練長度	16 epochs
單次訓練時間	約 3.5 小時

DOINN 的資料與架構對照使用共同前處理、最佳化器、batch size 與訓練長度；TopoNet 再依第 3.7 節加入對應架構與非對稱損失。官方跨架構權重則作基準參照，並非全部經本研究相同流程重訓。推論未使用 test-time augmentation 或 ensemble；同一評估內的連通元件與重疊判定設定一致。

4.1.2 資料池與測試資料

本研究使用 LithoBench 衍生資料池與真實電路切片，資料規模如表 9 所示。MetalSet 原始 16472 筆中有 50 筆檔案損壞，排除後可用 16422 筆；這 50 筆不列為模型失敗案例。表中的 MetalSet、MixSet、CtrlSet 與 LOOSet 數字為切分前資料池總量，實際訓練仍遵循第 3.6 節的來源切分。

表 9. 資料池與評估資料規模

資料	規模	用途／來源
MetalSet	16,422	有效資料池；LithoBench
StressMetal	659	篩選後壓力評估樣本
MixSet	19,622	物理擾動增強資料池
CtrlSet	19,622	等量重複抽樣資料池
LOOSet	18,822	移除指定擾動量的資料池
未擾動 cell	194	正常輸出與誤報評估
Ibex metal2	1,200	真實電路主測試切片
Ibex metal1／metal3	150／97	既有跨層切片
gcd metal1	150	既有跨設計切片

CtrlSet 為 MetalSet 加上 3200 個重複抽樣的樣本，其總樣本數與梯度步數與 MixSet 完全相同，但分佈不變，用於排除資料量的混淆。LOOSet 的擾動量僅包含 ± 1 、 ± 3 、 ± 6 ，用於測試模型對未見擾動量的泛化能力。

4.1.3 評估單位與統計彙整

表 10 將主要比較依輸出單位、參考規模與權重來源整理。資料池筆數、切片數、原始缺陷計數、匹配事件數與獨立訓練次數具有不同意義：例如 571 是橋接參考事件數，3 才是增強與 TopoNet 的訓練重複次數。不同測試表只有在光罩、事件定義與參考集合一致時才直接比較。

表 10. 核心評估的樣本單位與統計方式

評估	量測單位與規模	模型／權重	彙整方式
橋接完整配對	571 個參考事件	原版官方單次；增強、TopoNet 各 3 次	各次先算 P/R/F1，再平均與計算 SD
斷裂完整配對	383 個匹配參考事件	原版、增強、TopoNet 代表權重	單次 TP/FP/FN 與 P/R/F1
未擾動正常樣本	194 個 cell 切片	三個模型的代表權重	IoU、事件數、元件數與有誤報切片數
單次消融	共同壓力評估設定	dev_topo 完整模型及各消融版本	單次比較；不使用三次均值相減
Ibex metal2 主測試	1,200 切片；343 個參考事件	原始 ILT 光罩；三個模型代表權重	融合區事件配對；各版本獨立

註：P、R、SD 分別表示 precision、recall 與樣本標準差。參考事件來自本文所採物理模型；「正常」表示相對未擾動基準的評估條件。表 22 至表 24 列出前三項的完整數字，表 25 與表 32 分別呈現消融與 Ibex 主測試。

多種子指標先對每次訓練計算，再取算術平均與樣本標準差，四捨五入在彙整後進行。同一個參考事件可以出現在三次模型評估中，因此不將三次結果相加後宣稱樣本量增加為三倍。標準差描述已觀察到的訓練變異，不是信賴區間或正式顯著性檢定；同一切片內的事件也可能共享幾何特性。

資料增強對照固定 DOINN 架構；架構對照固定 MixSet；消融逐一移除機制。官方權重只作獨立參照，不混入本研究自訓的檢查點序列。各資料分割沿用第 3.6 節的來源切分，本文的事件抽取設定對同一比較中的所有模型相同。正常樣本與壓力樣本分開報告，以免缺陷篩選造成的事件比例掩蓋正常成像代價。

4.2 研究問題一 成像相似度與缺陷可靠度

本節檢查高全域 IoU 是否足以支持缺陷判斷。證據包含跨模型比較、依擾動量分組的事件 recall、低餘裕 hotspot 與首次失效幅度偏移。各項量測互補，並以事件位置配對作為缺陷是否被正確預測的依據。

4.2.1 全域成像指標與事件偵測

四個模擬器在缺陷層級的表現如表 11 所示。

表 11. 四個模擬器的缺陷層級表現

模型	全域 IoU	橋接 recall	斷裂 recall	消失 recall
DOINN	0.867	0.501	0.830	0.870
DAMOLitho	0.799	0.785	0.729	0.536
CFNOLitho	0.600	0.096	0.292	0.215
LithoGAN	0.226	0.112	0.078	0.873

(a) 同一 StressMetal sample 的多模型預測

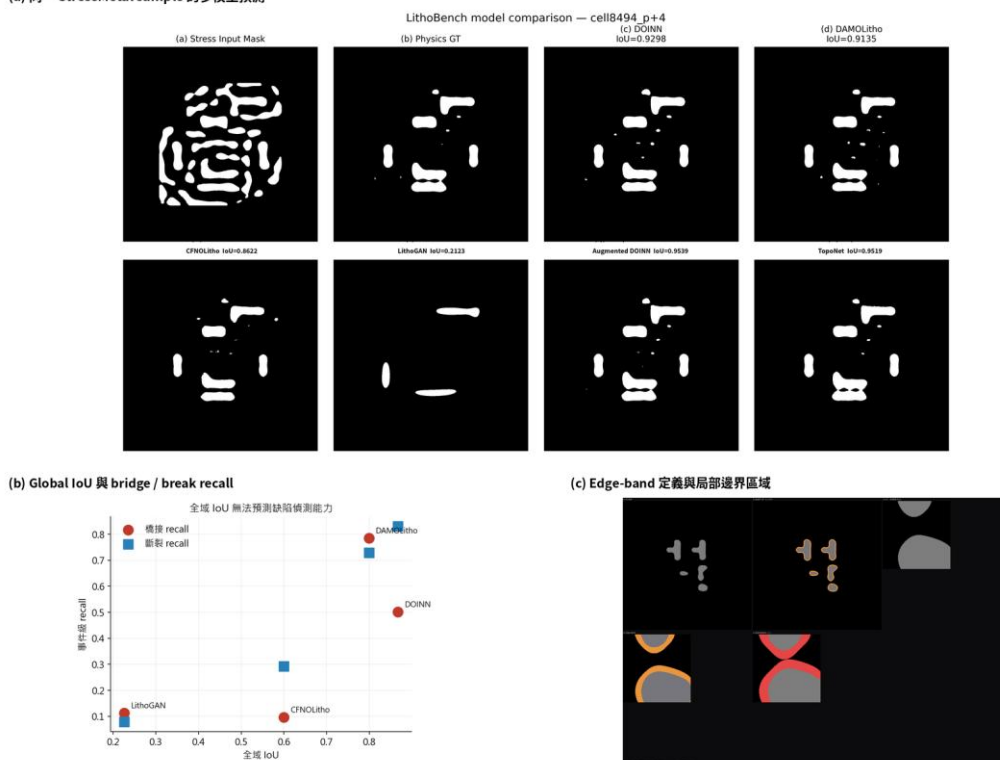


圖 6. 同一 StressMetal 樣本的多模型輸出、全域 IoU 與 bridge recall 關係，以及 edge-band 定義

DOINN 具有最高的全域 IoU，但其 bridge recall 仍低於 IoU 較低的 DAMOLitho，且各模型在不同缺陷類別上的相對表現並不一致。此結果顯示，全域像素相似度不足以單獨代表 lithography verification 所需的 defect-level reliability。圖 6(a)以同一個 +4 px StressMetal 樣本展示多模型輸出，其單張 IoU 僅作視覺範例；跨模型結論以表 11 的完整測試統計與圖 6(b)的關係為準。圖 6(c)則標示 edge-band 指標實際量測的窄帶區域。本研究進一步以 Spearman 等級相關分析各指標的預測力，結果如表 12 所示。

表 12. 全域／邊緣帶 IoU 與缺陷 recall 的 Spearman 等級相關 (n = 39 模型變體)

缺陷類型	全域 IoU	邊緣帶 IoU	較佳指標
橋接	+0.512	+0.771	邊緣帶
斷裂	+0.474	+0.351	全域
消失	+0.221	-0.013	全域

39 個觀測點為彙整後數值互異的模型變體，其中部分變體來自同一架構的不同 random seed，因此不能視為完全獨立的影像樣本。表 12 主要用於描述跨模型趨勢：bridge recall 與 edge-band IoU 的相關係數為 0.771，高於與 global IoU 的 0.512。本文不據此宣稱兩個相關係數之差已通過嚴格的獨立樣本顯著性檢

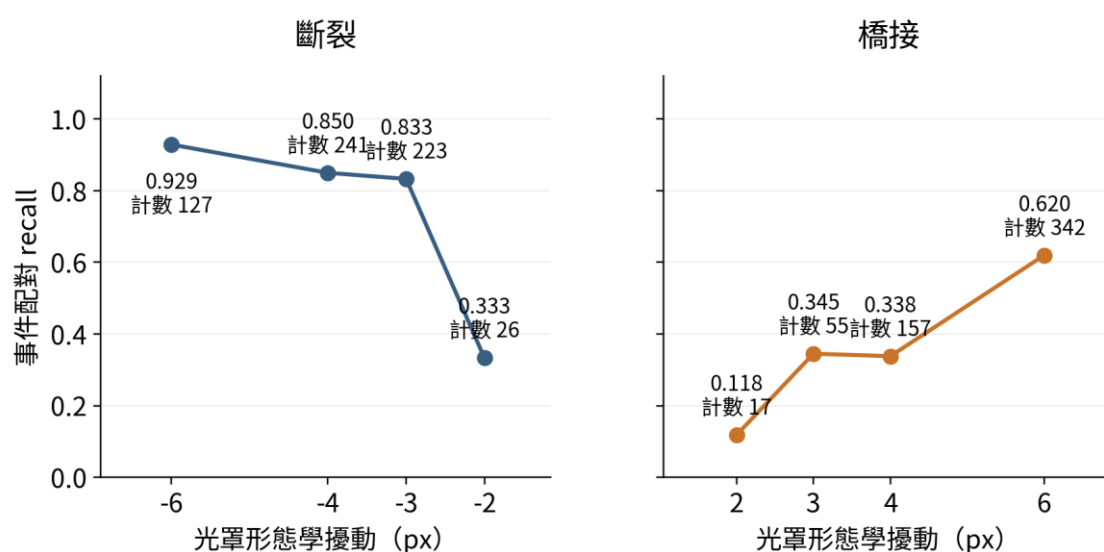
定。此差異與缺陷幾何特性一致：橋接主要發生於相鄰邊緣的接觸區，屬於帶內事件；斷裂與消失則與圖形存在性較相關。因而，缺陷層級評估仍需保留多種互補指標。

4.2.2 形態學擾動幅度與漏判

本研究依擾動量分組計算偵測率，結果如表 13 與圖 7 所示。

表 13. 依形態學擾動分組的事件計數與配對 recall

擾動 (px)	橋接計數	橋接 recall	原始斷裂計數	斷裂 recall
-6	0	—	127	0.929
-4	0	—	241	0.850
-3	0	—	223	0.833
-2	0	—	26	0.333
+2	17	0.118	0	—
+3	55	0.345	0	—
+4	157	0.338	0	—
+6	342	0.620	0	—



標註為各組原始事件計數；斷裂計數不作為事件配對 recall 的分母。

圖 7. 光罩形態學擾動下的事件配對 recall 與原始事件計數

本壓力集中的橋接出現在正向膨脹，斷裂出現在負向收縮；接近 nominal geometry 的低幅度組較難偵測。橋接 recall 在 +6 px 為 0.620，在 +2 px 為 0.118。表 3 的 MetalSet +2 px 與 +6.6% dose 僅依印出覆蓋率對齊，因此圖 7 橫軸使用「光罩形態學擾動 (px)」。負向四組的切片數依序為 167、156、84、

19，其中有斷裂的切片為 58、119、84、18；原始 brk_gt 計數則為 127、241、223、26。三種數量分別描述切片、受影響切片與原始事件，不能互換。原始斷裂計數總和 617 亦不作為第 4.4 節整體配對 recall 的分母；該評估以 TP+FN=383 個配對參考事件計算。

4.2.3 低餘裕位置的偵測率

本研究由壓力集整理 1080 個物理模擬所產生的缺陷位置，記錄座標、缺陷類型、原始間距與首次失效所需的擾動幅度，並依此「擾動餘裕」分為四個相對風險等級。各模型的偵測率如表 14 所示。圖 8 將相同結果依首次失效餘裕分組呈現，可直觀看出最低餘裕區間是基準模型最容易漏檢的位置。

表 14. 依風險等級的 hotspot 偵測率

風險等級	樣本數	DOINN	DAMOLitho	增強	TopoNet
極危（餘裕 2）	44	27.3%	56.8%	63.6%	79.5%
高（餘裕 3）	233	86.7%	86.3%	89.3%	95.7%
中（餘裕 4）	324	75.3%	61.1%	89.2%	92.0%
低（餘裕 6）	479	86.6%	82.3%	89.1%	90.6%

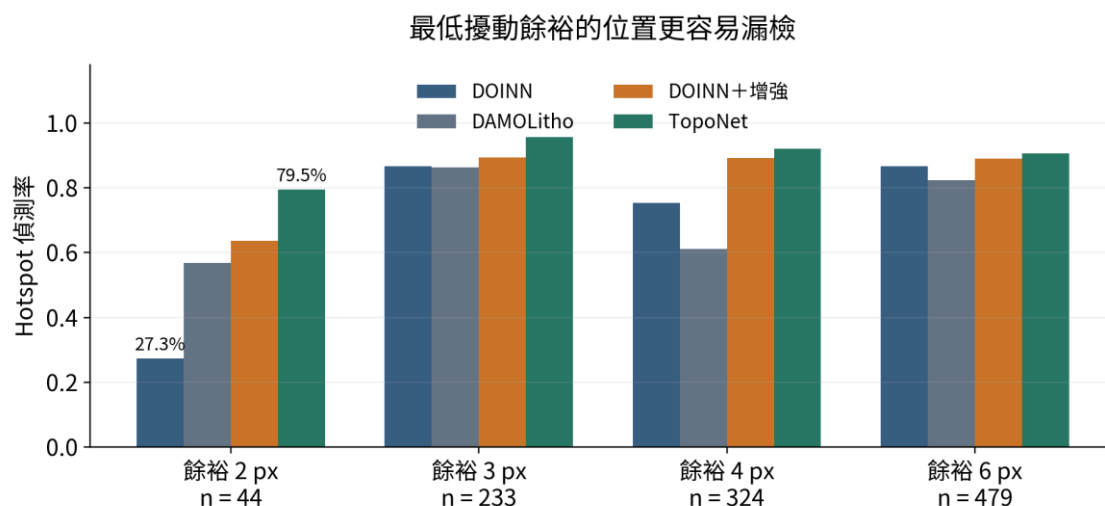


圖 8. 依首次失效擾動餘裕分級的 hotspot 偵測率

基準模型在其餘三個風險等級的偵測率約為 75% 至 87%，在最低擾動餘裕等級則降至 27.3%。此結果顯示盲區具有明顯的風險集中性，主要出現在接近 topology transition 的位置。

4.2.4 首次失效幅度的偏移

為量化模型是否把臨界位置誤判得過度安全，本研究以 stress set 中「首次

產生缺陷的形態學擾動幅度」作為離散餘裕，並定義 ΔPW 為模型與物理真值之首次失效擾動幅度差，如公式(7)所示。此處的 ΔPW 單位為擾動級距，而非實際曝光劑量。

$$\Delta PW = m_fail(pred) - m_fail(gt) \quad (7)$$

ΔPW 為有號量：正值表示模型需要更大的擾動才判定失效，即高估擾動餘裕；負值表示模型較早提出警報。它僅描述離散形態學掃描，不是經製程校準的曝光窗口。對掃描範圍內完全漏檢的位置，模型首次失效幅度並不存在有限觀測值，因此表 15 同時保留完全漏檢數；平均偏移量不單獨用來比較整體安全性。

表 15. 擾動餘裕的高估量

模型	完全漏檢	ΔPW (擾動級距)	高估比例
DOINN	20	+1.17	53.7%
DAMOLitho	12	+0.42	24.1%
CFNOLitho	75	+2.88	30.6%
LithoGAN	64	+0.43	13.0%

既有彙整結果中，DOINN 的平均偏移為 +1.17 個擾動級距，高估比例為 53.7%，呈現延後判定失效的傾向。由於完全漏檢與有限幅度偏移屬不同情況，本文以事件級 TP、FP、FN 與低餘裕 hotspot 偵測率作為主要可靠度證據， ΔPW 僅作輔助描述。

4.3 研究問題二 資料分佈與漏判改善

本節將模型端修改、資料幾何統計與訓練資料對照放在同一論證中。關鍵對照是 CtrlSet 與 MixSet：兩者資料池總量與每 epoch 梯度步數相同，僅前者採重複抽樣，後者加入物理重模擬的擾動樣本。

4.3.1 模型端修改的對照

在資料增強前，本研究先測試延長訓練、二值化閾值、間隙損失加權與尺度遞迴架構，結果如表 16 及圖 9 所示。前三組報告橋接 recall；尺度遞迴架構因成像品質明顯下降，只保留 mIoU，未計算 recall。不同實驗有各自對照，不合併為同一基準線。

表 16. 四組模型端修正方法的結果

嘗試	結果	範圍與判讀
延長訓練	自訓 1 ep : 0.515 ; 16 ep : 0.497 ; 官方 : 0.501	官方權重為獨立參照
預測二值化閾值	0.5 / 0.4 / 0.3 $\rightarrow$ 0.501 / 0.525 / 0.546	最大增加 4.5 個百分點
間隙損失加權	0.417 / 0.457 / 0.494	皆低於該組對照 0.578
尺度遞迴架構	mIoU 0.216 / 0.031	成像品質下降；未計算 recall

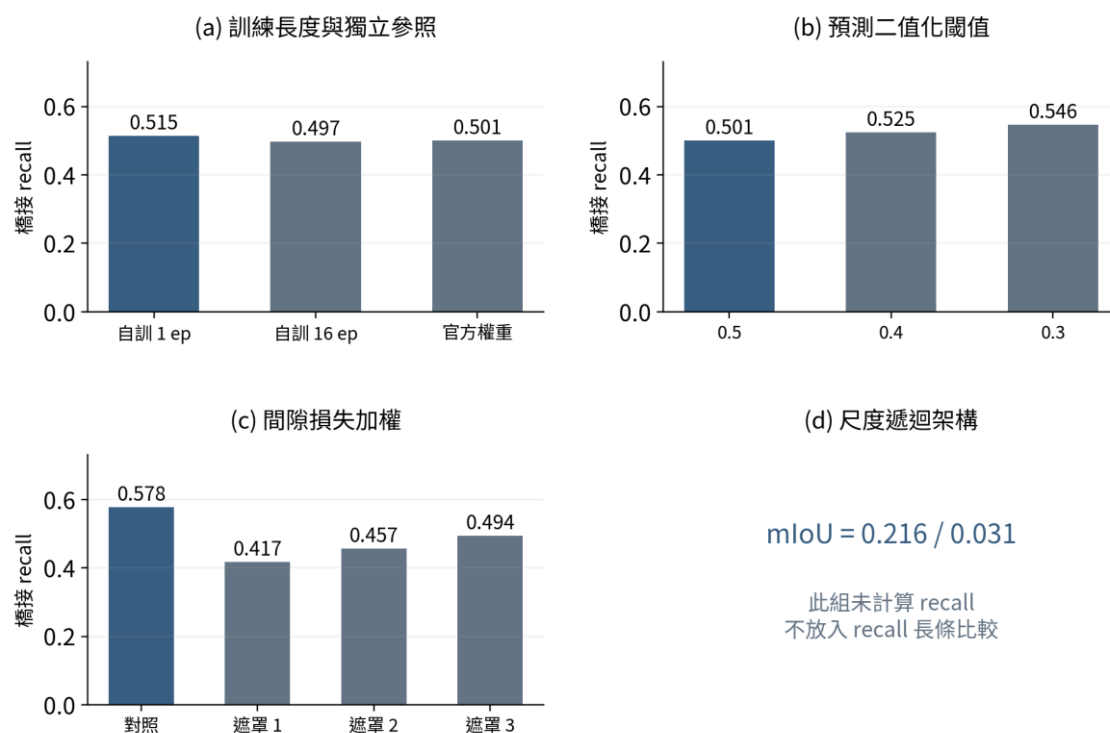


圖 9. 模型端修改的各自對照與尺度遞迴架構的 mIoU

本研究自訓 DOINN 的兩個檢查點為 1 epoch 與 16 epochs，橋接 recall 分別為 0.515 與 0.497；訓練延長期間全域 IoU 由 0.705 提升至 0.867，斷裂 recall 由 0.426 提升至 0.830。官方預訓練權重的橋接 recall 為 0.501，來自 LithoBench 另一次訓練，僅作獨立參照，並非本次訓練的第三個檢查點。結果顯示本設定下單純延長訓練未改善橋接漏判。間隙損失加權亦提供診斷證據。若問題主要來自 bridge sample 的局部權重不足，提高 gap 區域權重應能帶來改善；然而三種遮罩設計均未優於對照組，且 SRAF-like component 的誤印比值由 1.21 升至 1.31–1.42。結果顯示，僅調整 loss weighting 無法消除主要盲區。

4.3.2 訓練資料的近臨界覆蓋

模型端修改、資料量對照與樣本幾何統計共同支持：近臨界拓樸案例在訓練

分佈中的不足，是本研究 DOINN 漏判的重要因素。此判讀針對已測資料與設定，不表示模型容量、損失設計或其他訓練方式均已被排除。如表 2 所示，本研究抽樣量測的訓練真值中，最小間隙為 35 像素、第 1 百分位為 53 像素，1983 個元件配對中未觀察到小於 10 像素的近臨界 gap；相較之下，模型的邊緣位移誤差為 -0.694 ± 1.102 像素。兩者尺度差距顯示，原始訓練資料很少提供「數個像素誤差即可改變連通性」的監督訊號。圖 10 將訓練資料的最小 gap 與模型邊緣誤差放在同一尺度下比較，用以說明近臨界細縫缺乏訓練覆蓋的程度。

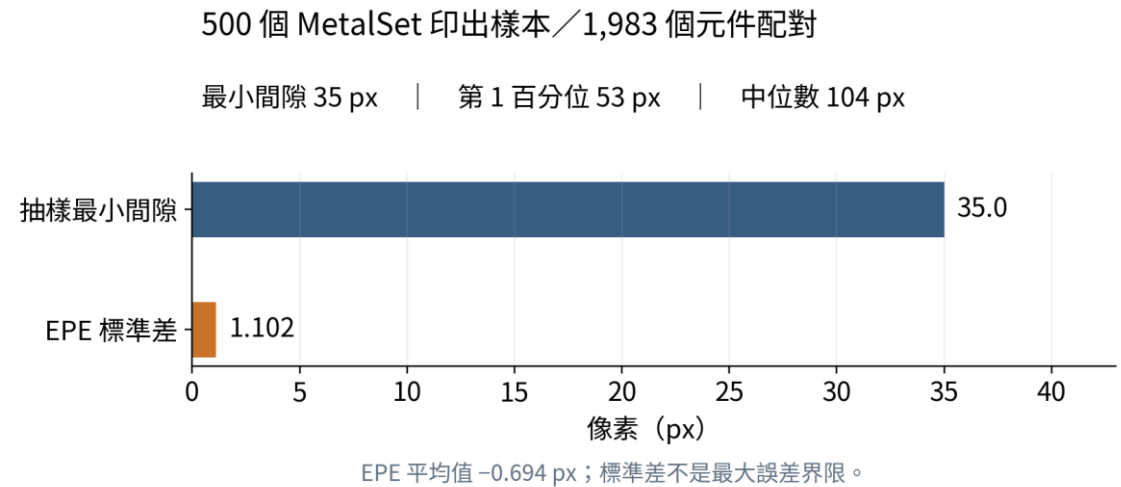


圖 10. 抽樣元件間隙與模型 EPE 標準差的尺度比較

正常樣本評估使用壓力集所涵蓋的 194 個 cell 之未擾動光罩。未擾動物理參考與自身比較時，橋接、斷裂及消失事件均為 0，這是事件抽取的自我一致性檢查，不能單獨證明其相對設計目標不存在所有製程缺陷。第 4.4 節進一步量測模型相對此基準產生的不存在事件，以補足只看缺陷 recall 的限制。

4.3.3 物理擾動增強

補上缺失的分佈後，結果如表 17 所示。

表 17. 物理擾動增強的效果

訓練資料	MetalSet 正常 IoU	壓力集 IoU	橋接 recall	參數
MetalSet	0.949	0.856	0.520 ± 0.052	211.6K
MixSet	0.973	0.911	0.8634 ± 0.0063	211.6K

註：IoU 為該組代表權重的單次結果；橋接 recall 為 MetalSet 四次、MixSet 三次訓練的平均 $\pm$ 樣本標準差。不同測試集合的正常 IoU 分開報告。

增強後的橋接 recall 為 0.8634 ± 0.0063 ，較既有重複訓練基準 0.520 提高

約 34.3 個百分點。表 17 的 MetalSet 正常測試中，IoU 由 0.949 提升至 0.973；此觀察限定於該批資料，不能推論所有正常圖形皆無品質取捨。第 4.4 節另以 194 個未擾動 cell 評估，則觀察到增強與 TopoNet 的 IoU 下降及誤報增加。

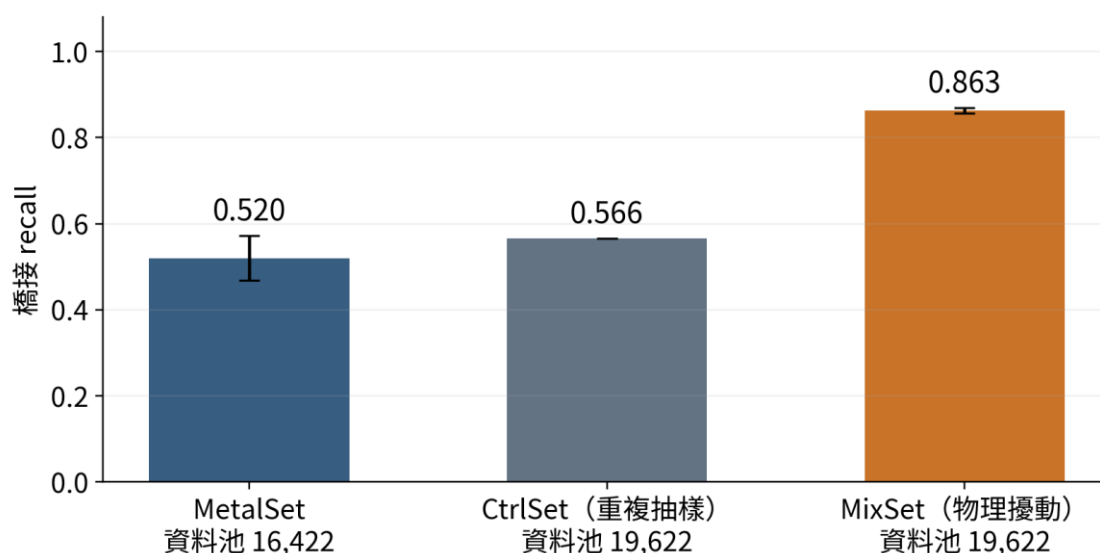
4.3.4 資料量與擾動涵蓋程度的控制

MixSet 比 MetalSet 多出 19% 的樣本，因此本研究建立 CtrlSet 作為對照：加入 3200 個重複抽樣的樣本，使總樣本數與每個 epoch 的梯度步數與 MixSet 完全一致，但分佈不變。結果如表 18 所示。圖 11 以相同訓練設定比較 MetalSet、等量擴充的 CtrlSet 與包含物理擾動樣本的 MixSet；只有改變資料分佈的 MixSet 帶來明顯 bridge recall 提升。

表 18. 資料池規模與每 epoch 梯度步數對照

訓練資料	資料池總量	步數／epoch	橋接 recall
MetalSet	16,422	3,695	0.520 ± 0.052
CtrlSet (重複抽樣)	19,622	4,415	0.566
MixSet (物理擾動)	19,622	4,415	0.8634 ± 0.0063

等量資料對照：缺陷分佈比重複抽樣更有幫助



誤差棒：MetalSet 4 次、MixSet 3 次訓練的標準差；CtrlSet 為單次結果。

圖 11. CtrlSet 與 MixSet 對照：改善主要來自資料分佈，而非單純增加樣本量

CtrlSet 的橋接 recall 為 0.566，較基準增加 0.046，但仍接近其重複訓練的變異尺度；MixSet 的三個隨機種子平均為 0.8634。此對照支持改善主要來自新

增缺陷分佈，而非僅增加重複樣本或梯度步數；單次 CtrlSet 結果不作統計顯著性結論。為檢查對特定擾動量的依賴，本研究建立 LOOSet：訓練擾動僅包含 ± 1 、 ± 3 、 ± 6 ，並於測試中分別統計 $\pm 2/\pm 4$ 組與其餘擾動組。MixSet 已包含 $\pm 2/\pm 4$ ，故表 19 直接標示實際擾動量，避免對所有列一律稱為「未見組」。

表 19. 擾動泛化測試

模型／訓練資料	訓練含 $\pm 2/\pm 4$	$\pm 2/\pm 4$ recall	其餘擾動 recall
DOINN／MixSet	是	0.753	0.902
DOINN／LOOSet	否	0.770	0.927
TopoNet／MixSet	是	0.902	0.970
TopoNet／LOOSet	否	0.833	0.950

DOINN 在 $\pm 2/\pm 4$ 組中，LOOSet 的 recall 為 0.770，略高於 MixSet 的 0.753；TopoNet 則由 MixSet 的 0.902 降為 LOOSet 的 0.833，相差 6.9 個百分點。故本次 DOINN 未觀察到移除這些擾動量造成下降，但 TopoNet 對訓練涵蓋程度較敏感。結果不能概括為「訓練是否見過擾動量不影響表現」，亦不能用單次差異證明某種訓練配置必然較佳。隨機種子的變異亦提供佐證，如表 20 所示。

表 20. 隨機種子變異

訓練配置	各個隨機種子橋接 recall	平均 $\pm$ 樣本標準差
MetalSet (4 次)	0.497／0.578／0.545／0.461	0.520 ± 0.052
MixSet (3 次)	0.8687／0.8564／0.8651	0.8634 ± 0.0063
TopoNet (3 次)	0.9492／0.9335／0.9545	0.9457 ± 0.0109

以各次訓練的完整精度事件統計計算，增強 DOINN 的橋接 recall 為 0.8634 ± 0.0063 ，TopoNet 為 0.9457 ± 0.0109 。MetalSet 四次訓練的基準為 0.520 ± 0.052 ；官方獨立權重則為單次 0.5009。本文分別標示重複訓練平均與單次結果，避免把不同權重來源視為同一實驗。

4.4 研究問題三 TopoNet 的改善與代價

在共同 MixSet 訓練條件下，本節評估 TopoNet 是否提供架構上的增益，並同時報告橋接、斷裂及正常樣本結果。橋接使用三個隨機種子，正常樣本與斷裂表使用代表權重；各自的統計單位見表 10。

4.4.1 TopoNet 架構比較

本研究先以視覺化檢查下採樣對細縫資訊的影響，如圖 12 所示。

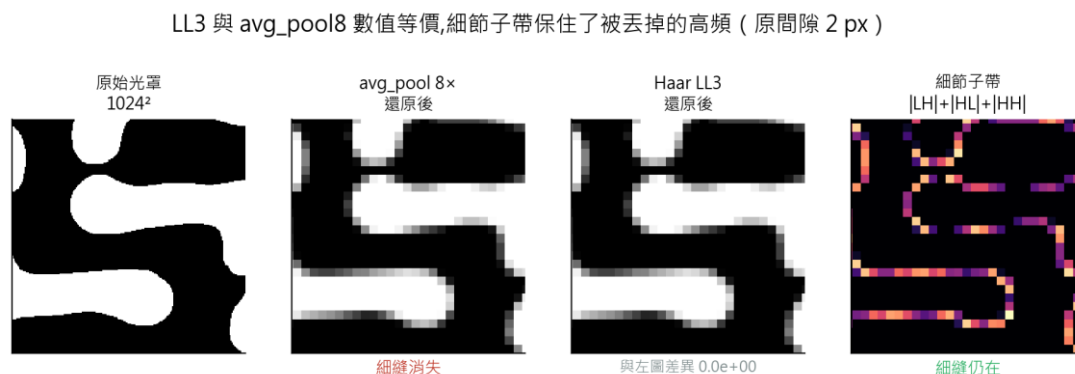


圖 12.8 倍平均池化與 Haar 小波對細縫高頻資訊的保留差異

圖中由左至右分別為原始光罩、8 倍平均池化後還原結果、Haar 小波低頻子帶還原結果，以及細節子帶的絕對值總和。低頻子帶與平均池化的最大差異僅在 10^{-7} 量級，兩者在數值上近似等價；Haar 分解額外提供的細節子帶則保留了平均池化所捨棄的高頻資訊。TopoNet 的整體表現如表 21 所示。

表 21. 架構比較 (皆於 MixSet 訓練)

模型	橋接 recall (3 次)	消失 recall	全域 IoU	參數量
DOINN+增強	0.8634 ± 0.0063	0.953	0.914	211.6K
WRD-Net (僅小波)	0.848 ± 0.039	0.960	0.915	213.3K
TopoNet	0.9457 ± 0.0109	0.966	0.9124	217.6K

註：橋接欄為三次訓練平均 $\pm$ 樣本標準差；消失與全域 IoU 為代表權重的單次評估。WRD-Net 僅小波對照的斷裂 recall 為 0.854；原版、增強與 TopoNet 的完整斷裂配對另列表 23。

TopoNet 的三個隨機種子平均橋接 recall 較增強 DOINN 提高 0.0823 (8.23 個百分點)，最低擾動餘裕 hotspot 偵測率亦由原版的 27.3% 提升至 79.5%。以下補上事件級 precision 與 F1，以及未擾動樣本上的誤報評估，使改善是否伴隨其他代價可被直接檢查。

4.4.2 橋接事件的完整配對結果

表 22 逐一列出 571 個橋接參考事件的配對結果。DOINN 原版為官方參照權重的單次評估；增強與 TopoNet 各列三個隨機種子。此表的 0.5009 不取代原始資料四次訓練平均 0.520，兩者的權重來源與彙整方式不同。

表 22. 橋接事件配對結果（每列 571 個參考事件）

模型／種子	TP	FP	FN	Precision	Recall	F1
DOINN 原版	286	0	285	1.0000	0.5009	0.6674
增強 seed 1	496	13	75	0.9745	0.8687	0.9185
增強 seed 2	489	2	82	0.9959	0.8564	0.9209
增強 seed 3	494	11	77	0.9782	0.8651	0.9182
TopoNet seed 1	542	36	29	0.9377	0.9492	0.9434
TopoNet seed 2	533	23	38	0.9586	0.9335	0.9459
TopoNet seed 3	545	35	26	0.9397	0.9545	0.9470

註：每列 TP+FN = 571；原版無 FP 表示本次預測事件皆配對成功，不表示已找到所有事件。

由未四捨五入的各次指標計算，增強 DOINN 的 precision、recall、F1 平均 ± 樣本標準差依序為 0.9829 ± 0.0115 、 0.8634 ± 0.0063 、 0.9192 ± 0.0015 ；TopoNet 為 0.9453 ± 0.0116 、 0.9457 ± 0.0109 、 0.9454 ± 0.0018 。TopoNet 的 recall 較增強高 8.23 個百分點、F1 高 2.62 個百分點，但 precision 低 3.75 個百分點，說明漏檢改善伴隨較多誤報。

4.4.3 斷裂事件的完整配對結果

表 23. 斷裂事件配對結果（每列 383 個參考事件）

模型	TP	FP	FN	Precision	Recall	F1
DOINN 原版	318	72	65	0.8154	0.8303	0.8228
DOINN+增強	314	35	69	0.8997	0.8198	0.8579
TopoNet	335	35	48	0.9054	0.8747	0.8898

註：本表為代表權重的單次評估；每列 TP+FN = 383。TopoNet recall = $335/(335+48) = 0.8747$ ，四捨五入為 0.875。

增強 DOINN 的斷裂 recall 為 0.8198，略低於原版 0.8303，但 precision 由 0.8154 提高至 0.8997，F1 由 0.8228 提高至 0.8579。TopoNet 的 precision、recall 與 F1 分別為 0.9054、0.8747、0.8898，皆高於本表另外兩者。因此不能概括成「資料增強讓每種缺陷的 recall 都增加」，應分別說明各項取捨。

4.4.4 未擾動樣本的成像品質與誤報

另取壓力集來源涵蓋的 194 個 cell 之未擾動光罩，以相同物理參考與事件抽取流程檢查模型是否產生不存在的事件。物理參考與自身比較時，橋接、斷裂

及消失皆為 0；表 24 報告相對此未擾動基準的模型輸出。這批樣本與表 17 的 MetalSet 正常測試不同，兩組 IoU 不混合彙整。

表 24. 194 個未擾動 cell 的成像品質與誤報

模型	IoU	誤報橋接	誤報斷裂	誤報消失	新增元件	有誤報切片
DOINN 原版	0.9216	0	1	0	6	1 / 194
DOINN + 增強	0.8935	1	3	0	17	4 / 194
TopoNet	0.8207	2	14	0	471	14 / 194

註：橋接、斷裂、消失欄為事件數；新增元件另以元件數計，有誤報切片依評估輸出的切片旗標統計。新增元件不直接換算為短路或功能失效。

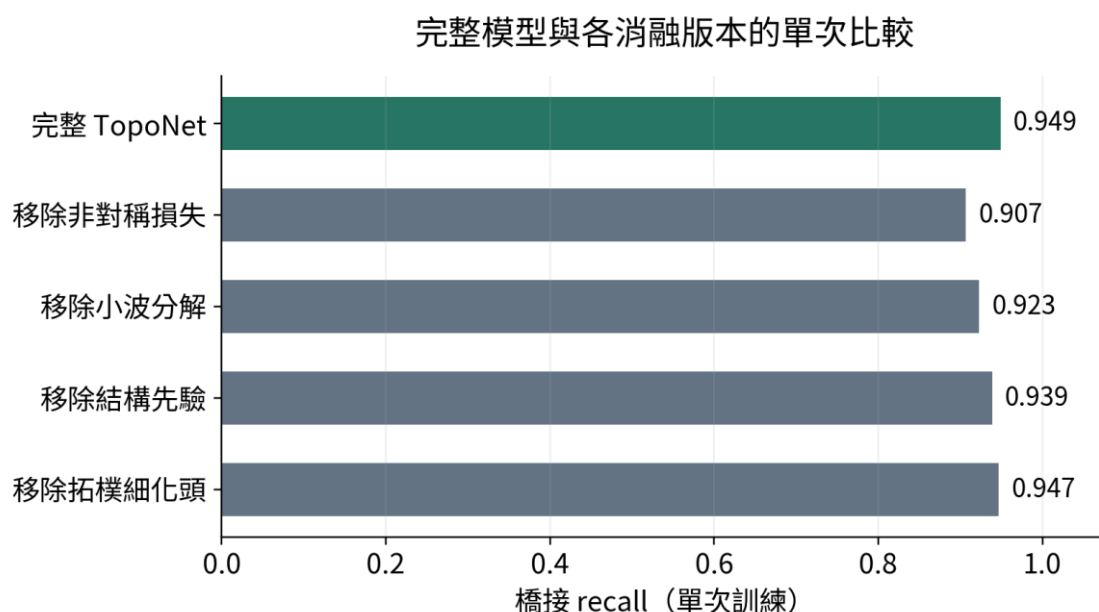
原版、增強與 TopoNet 的有誤報切片比例分別為 0.52%、2.06%、7.22%，而 IoU 依序為 0.9216、0.8935、0.8207。TopoNet 另產生 471 個新增元件，顯示正常輸出的細小結構需進一步抑制。結合表 22 與表 23，本文結論是壓力集事件品質改善，但尚有正常樣本保真度與誤報代價；目前結果不足以支持直接、全面替代物理模擬。

4.4.5 單次消融分析

為觀察各機制的作用，本節逐一移除單一機制並重新訓練，結果如表 25 及圖 13 所示。完整 TopoNet 使用 dev_topo 的單次結果（橋接 0.949、斷裂 0.875、全域 IoU 0.9124、邊緣帶 IoU 0.6379），與各單次消融列比較；主結果的三個隨機種子平均 0.9457 ± 0.0109 不作此表的相減基準。

表 25. 消融實驗結果

版本	橋接 recall	斷裂 recall	全域 IoU	邊緣帶 IoU
完整 TopoNet (單次)	0.949	0.875	0.9124	0.6379
移除非對稱損失	0.907	0.896	0.924	0.690
移除小波分解	0.923	0.875	0.897	0.567
移除結構先驗	0.939	0.843	0.906	0.637
移除拓撲細化頭	0.947	0.877	0.907	0.623



完整模型採 dev_topo : 0.949 ; 三種種子平均 0.9457 另列主結果。

圖 13. TopoNet 各機制的單一 seed 消融結果

以完整模型的單次橋接 recall 0.949 為基準，移除非對稱損失、小波、結構先驗與細化頭後，分別下降 0.042、0.026、0.010 與 0.002。移除非對稱損失時，全域 IoU 升至 0.924、邊緣帶 IoU 升至 0.690，但橋接 recall 降至 0.907，顯示像素保真度與橋接敏感度可能有不同最佳化方向。細化頭的單次差異僅為 0.002，尚不足以確認穩定增益或宣稱沒有貢獻。消融結果提供結構趨勢，未作多種種子的模組效應估計。

4.5 Ibex 延伸驗證與電路連線風險

本節保留電路設計的延伸驗證：檢查 LithoBench 上完成訓練的模型能否在不同版圖中維持改善方向，並利用 net 標籤區分幾何融合與跨網路短路。這些結果補充核心研究的遷移證據，不與 LithoBench 的事件分母合併。

4.5.1 版圖與光罩版本

本節起將第一階段訓練完成的模型直接套用於真實電路的版圖，過程中不進行任何重新訓練或微調。受測電路的規格如表 26 所示。

表 26. 受測電路

電路	標準元件數	連線數	晶片尺寸	DRC
Ibex RISC-V 處理器	13199 (DEF components)	17220	$239.8 \times 239.8 \mu\text{m}$	0 violations
gcd	986	759	$35.8 \times 35.8 \mu\text{m}$	0 violations



圖 14. OpenROAD 產生之 Ibex RISC-V routed layout 全晶片視圖

OpenROAD detailed-routing report 對 Ibex 回報 0 個 routing DRC violation、0 個 antenna violation，且 routing completion 為 100%。因此，可確認該設計在本次 Nangate45/OpenROAD routing rule check 下為 clean。本研究未執行 Calibre、KLayout 或 Magic 的完整 foundry sign-off DRC，也未完成 LVS。圖 14 提供同一 Ibex routed layout 的全晶片視圖，作為後續隨機切片與 net-level electrical-short 分析的版圖來源。製程擾動前先檢查光罩的 nominal topology。原始梯度 ILT 在 1200 個 Ibex metal2 切片中產生 8 處 nominal electrical short，本文將其視為實驗光罩生成 artifact，從該版本的製程擾動統計排除。另以 gap protection 建立 topology-constrained 控制，使 nominal short 降為 0。後續表格各自標明主測試、控制與跨層補充的用途；不同光罩版本不合併為同一批事件。topology-constrained ILT 控制在 1200 個 Ibex metal2 切片的 +1/+2 px 壓力下共有 200 個物理 electrical-short event，原版、增強與 TopoNet 的

recall 分別為 0.770、0.910、0.910，F1 分別為 0.830、0.873、0.892。另一次 ilt3 彙整回報總缺陷 200 處，其中 +2 px 子組含 176 處事件，三模型 recall 為 0.761、0.909、0.915；200 為整批計數，176 為子組計數，兩者不互換。上述控制與補充分組不取代原始 ILT 主表，也不與第 4.5.7 節跨層表的統計直接合併。

4.5.2 劑量方向的失效餘裕

在排除 8 個 nominal mask artifact 後，本研究沿劑量軸以 10 個離散操作點掃描 1200 個 Ibex metal2 切片，共得到 740 處製程擾動下的潛在 electrical-short site。為便於比較，本研究以 5%、11% 與 20% 作為分析分級門檻；這些門檻是本文的相對風險分組，不宣稱等同於特定 foundry 的 sign-off specification。結果如表 27 所示。

表 27. Ibex 版圖的劑量方向製程餘裕分佈 (740 處)

餘裕等級	處數	比例	說明
極低餘裕 (<5%)	1	0.1%	極小劑量偏移即失效
低餘裕 (5–11%)	23	3.1%	本研究設定的低餘裕 警示區間
中等餘裕 (>11–20%)	71	9.6%	需較大劑量偏移才失效
高餘裕 (>20%)	645	87.2%	在本次掃描範圍內餘裕較高

表 28. 劑量方向製程餘裕最小的位置

連線 A	連線 B	餘裕	所屬單元
rf_reg[399]	net2384	3.0%	暫存器檔案
09158	_09768_	5.0%	內部節點
alu_i.adder_in_b[12]	net2085	7.0%	ALU 加法器輸入
08958	rf_reg[847]	7.0%	暫存器檔案
03744	net1763	7.0%	內部節點

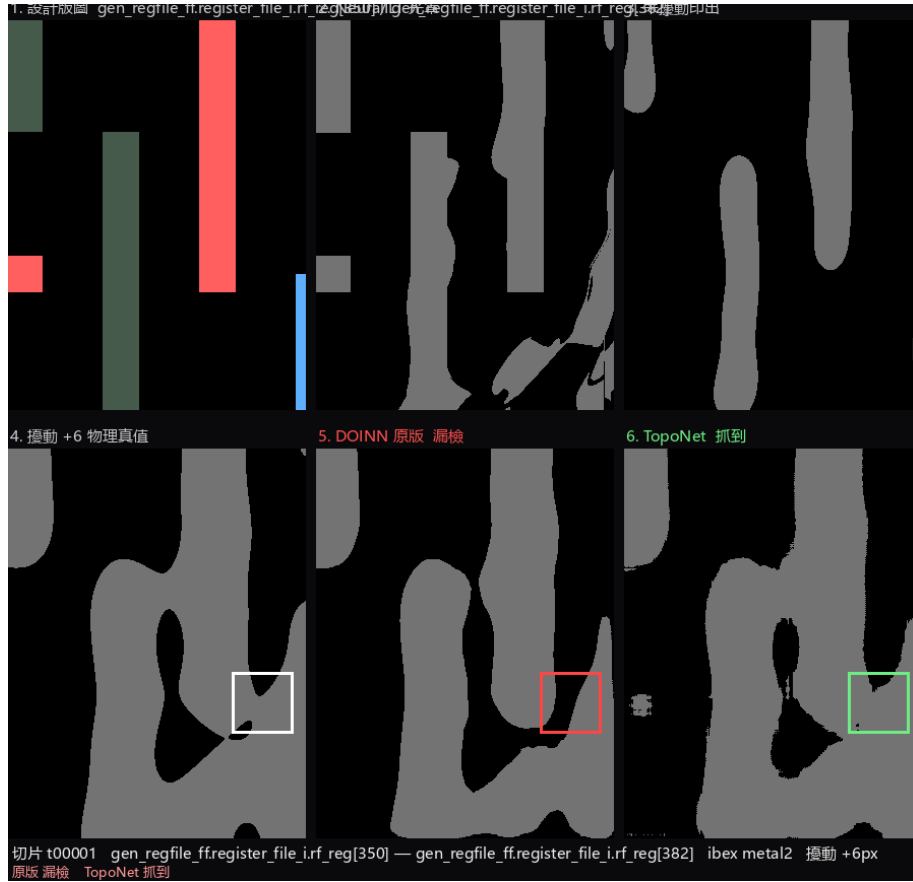


圖 15. Ibex metal2 代表性 electrical-short 案例：design/net、物理真值與三種神經模擬器輸出

表 28 所列部分低餘裕事件涉及暫存器檔案與 ALU 相關訊號。例如 rf_reg[399] 與 adder_in_b[12] 所在位置分別在 3.0% 與 7.0% 劑量偏移下出現 short。依 net 名稱與邏輯位置推測，這些事件可能影響暫存器資料或算術輸入；由於本研究尚未執行 Ibex gate-level fault injection 或 instruction-level simulation，本文僅將其描述為潛在功能風險。圖 15 以一個 metal2 代表案例對照 design/net、物理真值與多個神經模擬器輸出，顯示不同模型對同一跨網路橋接位置可能產生不同判定。

4.5.3 失效的幾何選擇性

為檢驗失效是否具有幾何選擇性，本研究量測每個失效位置在設計版圖上的原始間距，結果如表 29 所示。

表 29. 依設計間距分組的失效率

間距 (nm)	配對數	失效數	失效率
< 140	1044	91	8.7%
140–200	373	39	10.5%
200–280	455	44	9.7%
280–400	1264	22	1.7%
400–600	1301	10	0.8%
600–800	1081	1	0.1%

失效率未呈現嚴格單調關係：<140、140–200、200–280 奈米三組分別為 8.7%、10.5% 與 9.7%；當原始間距增加至 280 奈米以上後，失效率則明顯降至 1.7% 以下。全部配對的間距中位數為 346 奈米，失效配對的中位數為 142 奈米，約為前者的 40.9%。此外，84.1% 的失效發生在 280 奈米以內的配對，而此類配對僅佔全部樣本的 33.9%。因此，280 奈米可視為本資料集中的經驗性轉折尺度，且失效具有明顯幾何選擇性；該數值不應解讀為 Nangate45 的固定設計規則或特定 routing-track 公式。本研究另追蹤 32 個失效配對在離散劑量點上的狀態；在所測範圍內，事件一旦出現便未於更高劑量點恢復。此單調性與正向劑量壓力的預期方向一致，並降低「結果主要由隨機數值波動造成」的可能性，但不等同於完整製程校準。

4.5.4 單一間距門檻的限制

由於失效與原始間距具有明顯關聯，本節進一步評估單一 spacing threshold 是否足以篩選高風險位置。本研究將簡化幾何門檻由 65 奈米逐步放寬至 400 奈米，結果如表 30 所示。Nangate45 metal2 實際採用 context-dependent SPACINGTABLE，因此此實驗僅作為 single-threshold baseline，不代表完整 DRC deck。

表 30. 單一幾何門檻的涵蓋率與誤報占比（86 個參考事件）

門檻 nm	標記數	TP	FP	FN	涵蓋率	誤報占比
65	0	0	0	86	0.00%	—
80	10	1	9	85	1.16%	90.00%
100	28	4	24	82	4.65%	85.71%
130	2,980	35	2,945	51	40.70%	98.83%
160	3,804	56	3,748	30	65.12%	98.53%
200	4,302	60	4,242	26	69.77%	98.61%
280	5,680	78	5,602	8	90.70%	98.63%
400	9,504	83	9,421	3	96.51%	99.13%

幾何門檻的配對統計以 86 個物理參考事件為共同分母。130 奈米門檻涵蓋 35/86 (40.70%)，2980 個標記中有 2945 個未與參考事件配對，誤報占比為 98.83%；放寬至 400 奈米時涵蓋 83/86 (96.51%)，仍漏掉 3 個事件，誤報占比為 99.13%。此比例為 $FP/(TP+FP)$ ，不是需要 TN 的 FPR。65 奈米組沒有任何標記，故其誤報占比未定義。結果顯示已測單一間距門檻存在涵蓋率與誤標的明顯取捨，不能據此推廣為完整 DRC 或所有幾何方法都無效。

4.5.5 免重訓遷移結果

本節檢驗第一階段的修復方法是否具備免重訓遷移能力。模型完成 benchmark 訓練後，不針對 Ibex 重新訓練或微調，直接套用至真實 routed-layout 切片，結果如表 31 所示。

表 31. LithoBench 與 Ibex 原始 ILT 主測試的 recall

模型	LithoBench (多次平均)	Ibex metal2 (代表權重)
DOINN 原版	0.520	0.726
DOINN+增強	0.8634	0.892
TopoNet	0.9457	0.915

三個模型在 benchmark 與 Ibex 主測試中維持相同改善方向，但兩者的絕對數值不宜直接比較。LithoBench 欄的原版為既有四次訓練平均，增強與 TopoNet 為三次平均；Ibex 欄為所用代表權重的單次評估。前者使用篩選出的拓模壓力案例，後者由 routed DEF 隨機抽取 1200 個 2048 nm × 2048 nm 切片，再依所測條件統計事件。本研究未進行全晶片窮舉掃描，亦不將此結果解讀為全晶片故障率。

4.5.6 事件指標與二值化閾值

可靠度同時包含不漏檢 (recall) 與不誤報 (precision)。本研究由模型預測輪廓抽取 electrical-short site，再與相同光罩條件下的物理真值進行 event matching，以統計 TP、FP 與 FN。表 32 為原始 ILT 比較流程下 1200 個 Ibex metal2 切片的結果；第 4.5.1 節的 topology-constrained ILT 則作為獨立控制。

表 32. precision 與 recall (1200 切片)

模型	TP	FP	FN	precision	recall	F1
DOINN 原版	249	29	94	0.896	0.726	0.802
+ 增強	306	79	37	0.795	0.892	0.841
TopoNet	314	75	29	0.807	0.915	0.858

原版模擬器的 precision 最高 (0.896) 而 recall 最低 (0.726)。為確認此差異是否僅為判定閾值的選擇所致，本研究掃描閾值 0.3 至 0.7，結果如表 33 所示。

表 33. 判定閾值掃描 (200 切片)

閾值	原版 P	原版 R	原版 F1	TopoNet P	TopoNet R	TopoNet F1
0.3	0.815	0.830	0.822	0.586	0.962	0.729
0.4	0.857	0.792	0.824	0.646	0.962	0.773
0.5	0.857	0.792	0.824	0.729	0.962	0.829
0.6	0.857	0.792	0.824	0.803	0.925	0.860
0.7	0.891	0.774	0.828	0.842	0.906	0.873

在這批 200 個切片的閾值掃描中，原版模型 F1 約介於 0.822 與 0.828；TopoNet 則由 0.729 上升至 0.873，於閾值 0.7 時取得 precision 0.842、recall 0.906。這表示 TopoNet 的誤報與漏檢可隨閾值調整，而原版在已測範圍內的改善有限。此表僅描述測試條件的敏感度，未以獨立驗證集選擇部署閾值，因此不宣稱 0.7 是對所有資料最佳的設定。

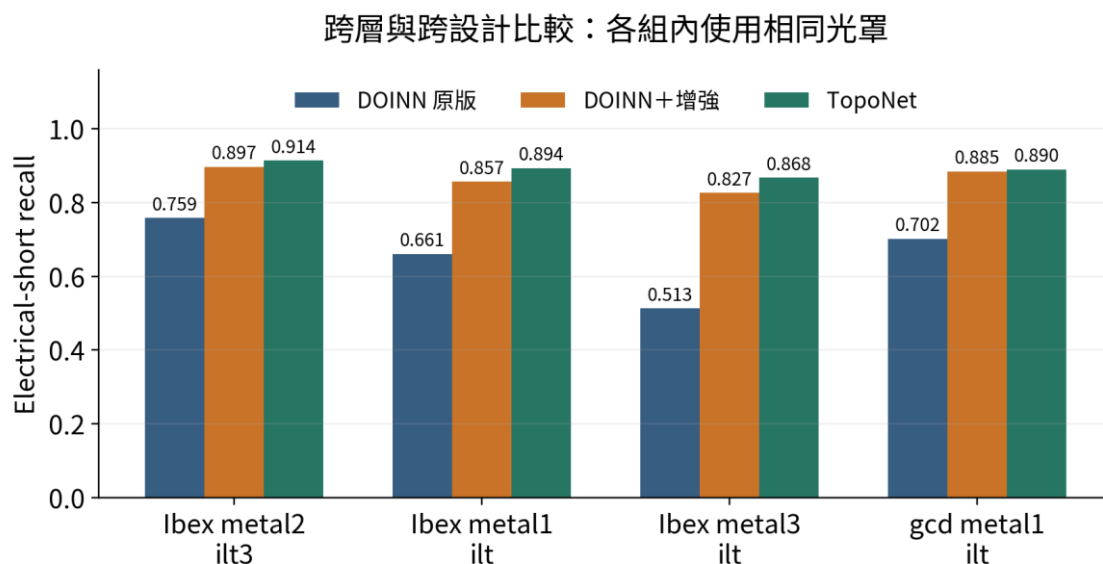
4.5.7 跨層與跨設計比較

跨層評估涵蓋 Ibex metal1/metal2/metal3 與 gcd metal1，結果如表 34。Ibex metal2 使用 ilt3，其餘三組使用 ilt；同一列的三個模型使用相同光罩版本。不同列的設計、金屬層及光罩版本不完全相同，因此本文比較各組內的改善方向，不將跨組 recall 高低歸因於單一因素。

表 34. 跨層與跨設計的 recall (標示光罩版本)

資料集	光罩版本	原版	增強	TopoNet
Ibex metal2	ilt3	0.759	0.897	0.914
Ibex metal1	ilt	0.661	0.857	0.894
Ibex metal3	ilt	0.513	0.827	0.868
gcd metal1	ilt	0.702	0.885	0.890

四組原版 recall 介於 0.513 與 0.759，增強後介於 0.827 與 0.897，TopoNet 介於 0.868 與 0.914。各組皆為 TopoNet > 增強 > 原版，支持改善方向在已測資料上的一致性；此排序不等同於各差異均通過統計顯著性檢定。圖 16 對應相同四組數據。



ilt / ilt3 為實驗光罩版本；跨資料集絕對數值不作同條件比較。

圖 16. 跨層與跨設計的 electrical-short recall 比較

metal3 的原版 recall 最低 (0.513)，但目前未將圖形密度、光罩品質與缺陷幾何分別控制，因此不由此結果單獨推論其失效機制。後續可在共同光罩生成與事件分層條件下，分析不同層別的難度來源。

4.5.8 布林抽象下的功能風險

在標準元件分析中，共取得 1246 個 electrical-short event。其連線類型統計使用全部 1246 個事件；其中 111 個事件因功能標註資訊或資料品質不足，無法可靠進行布林功能分類，因此在功能失效型態分析中排除。表 35 統計其餘 1135 個可有效分類事件。

表 35. 可有效分類之功能失效型態（1135 事件）

失效型態	事件數	比例
降維	1035	91.2%
輸入釘死	38	3.3%
輸出回授	35	3.1%
輸出釘死	18	1.6%
退化	9	0.8%
合計	1135	100.0%

以全部 1246 個事件的連線類型而言，1190 個（95.5%）為 signal-signal bridge，56 個（4.5%）涉及 signal-power；本次資料中未觀察到 power-power bridge。此統計僅描述事件組成，實際電流與失效嚴重度仍需橋接電阻與電路狀態等更完整的電性模型判定。在可有效分類的 1135 個事件中，邏輯降維佔多數；於本文的布林抽象下，這類事件表現為邏輯函數自由度下降或輸入關係被改寫，說明幾何 short 可能對應到不易由版圖外觀直接判斷的功能錯誤。表 36 進一步列出四類主要功能失效型態的模型偵測率；「退化」類僅 9 個事件，未納入此表的主要型態比較。

表 36. 依功能失效型態分類的偵測率

失效型態	事件數	DOINN	增強	TopoNet
輸出釘死	18	0.778	1.000	1.000
輸出回授	35	0.971	0.971	1.000
降維	1035	0.958	0.974	0.996
輸入釘死	38	1.000	1.000	1.000

基準模型在「輸出釘死」類型上的 recall 為 0.778。以 AND4_X1 的案例為例，在本研究的布林 short abstraction 下，輸入與接地連結會使輸出邏輯被限制；該事件在物理真值中存在，但基準神經模擬器未預測到對應 bridge。

4.6 壓力代理與物理參考的有效範圍

形態學擾動的用途在於建立可輸入既有 mask-only 模型的壓力案例。本節以覆蓋率、事件位置與印出閾值分析其有效範圍，使第 4.2 至 4.4 節的結論與實際量測條件相連。

4.6.1 形態學擾動與直接劑量的差異

本研究以形態學擾動作為神經模擬器可接受的 process-stress proxy。為檢查此代理與 direct-dose sweep 的關係，本節先以印出覆蓋率對齊兩者的操作點，再比較所產生的缺陷位置，結果如表 37 所示。

表 37. 形態學製程壓力與直接劑量掃描的比較

擾動量	覆蓋率對齊劑量	擾動覆蓋率	劑量覆蓋率	位置 Jaccard
+1 px	1.070	11.56%	11.42%	0.353
+2 px	1.165	17.95%	17.60%	0.140
+3 px	1.260	25.63%	25.26%	0.166

兩種變異方式的整體印出覆蓋率差異小於 0.4 個百分點，但缺陷位置僅部分重合。形態學直接改變光罩幾何，direct dose 則改變成像強度與閾值的交互作用，因此不能以覆蓋率相近推論局部失效等效。所採 ILT 光罩的線端 EPE（-7.99 奈米）亦大於側壁（-1.01 奈米），顯示位置反應並不均一。現有受測神經模擬器只接受光罩、不接受 dose 參數，因此本文採形態學壓力作為可輸入的測試方式；第 4.6.3 節另量測它與 dose failure 的涵蓋關係，界定此代理的限制。本研究亦檢視光罩品質的影響。不同資料集的 ILT 印出 IoU 約介於 0.66 至 0.78，且 cross-layer/cross-design 實驗中的改善方向保持一致，顯示目前結論在已測設定下未被單一批次光罩品質所主導。mask-generation quality 仍可能影響事件數量與難度，因此本文將其列為研究限制。

4.6.2 物理印出閾值的敏感度

物理參考使用 $Z = \text{sigmoid}[s \times (\text{aerial} - t)]$ 並以 $Z > 0.5$ 二值化，其中 s 為正陡度、 t 為印出閾值。由 $\text{sigmoid}(x) > 0.5 \Leftrightarrow x > 0$ ，可知二值輸出等價於 $\text{aerial} > t$ ；僅改變正陡度 s 不會改變輪廓。實際檢查 $s = 25$ 、 50 、 100 的三張影像時，二值結果差異皆為 0。因此陡度掃描不作為獨立穩健性證據，表 38 僅保留會改變二值輪廓的閾值設定。

表 38. 物理印出閾值變動下的 recall (500 個切片)

物理參考設定	參考事件數	原版	增強	TopoNet
基準閾值	82	0.744	0.890	0.902
閾值 $\times 0.85$	213	0.286	0.347	0.366
閾值 $\times 1.15$	19	0.947	1.000	1.000

基準與閾值 $\times 0.85$ 兩組皆呈現 TopoNet > 增強 > 原版；閾值 $\times 1.15$ 僅有 19 個參考事件，增強與 TopoNet 同為 1.000，故僅描述結果，不作強泛化結論。閾值降低時參考事件由 82 增至 213，三模型 recall 同時下降。由於神經模型未輸入閾值條件、也未針對此設定重訓，這反映物理參考條件改變時的敏感度；尚不能單憑排序不變宣稱已驗證完整光阻模型或製程穩健性。

4.6.3 兩種壓力模式的事件涵蓋

第 4.6.1 節顯示 morphology 與 direct dose 產生的缺陷位置僅部分重合。本節改以 failure coverage 量化兩者關係，即統計劑量掃描所找到的缺陷中有多少同時被形態學擾動涵蓋，結果如表 39 所示。

表 39. 兩種變異方式的涵蓋關係 (300 個切片)

操作點	擾動找到	劑量找到	交集	劑量被涵蓋的比例
+1 px (覆蓋率對齊 +7% dose)	7	3	3	100.0%
+2 px (覆蓋率對齊 +16.5% dose)	51	18	15	83.3%

形態學擾動對 direct-dose failure 具有高涵蓋率：低幅度主評估操作點 +1 px 涵蓋 3/3 (100%)，+2 px 涵蓋 15/18 (83.3%)。+1 px 在本批資料中恰好涵蓋全部 direct-dose failure；+2 px 則未形成數學上的 superset。形態學擾動另外產生的事件 (+1 px 時 4 處、+2 px 時 36 處) 在其自身物理重模擬真值下確實構成拓撲失效，但不應直接視為 direct-dose failure，可理解為對 mask geometry / critical-dimension 類變動較敏感的壓力案例。上述有限樣本顯示，形態學壓力可補充對幾何變動敏感的案例，但不是 direct-dose failure 的完整超集合。它額外產生的事件可能增加後續物理複核需求，而漏掉的 dose failure 仍需獨立條件評估。圖 17(a-d)為兩種模式共有的案例，圖 17(e-h)為 morphology-only 案例，用以呈現代理的用途與限制。

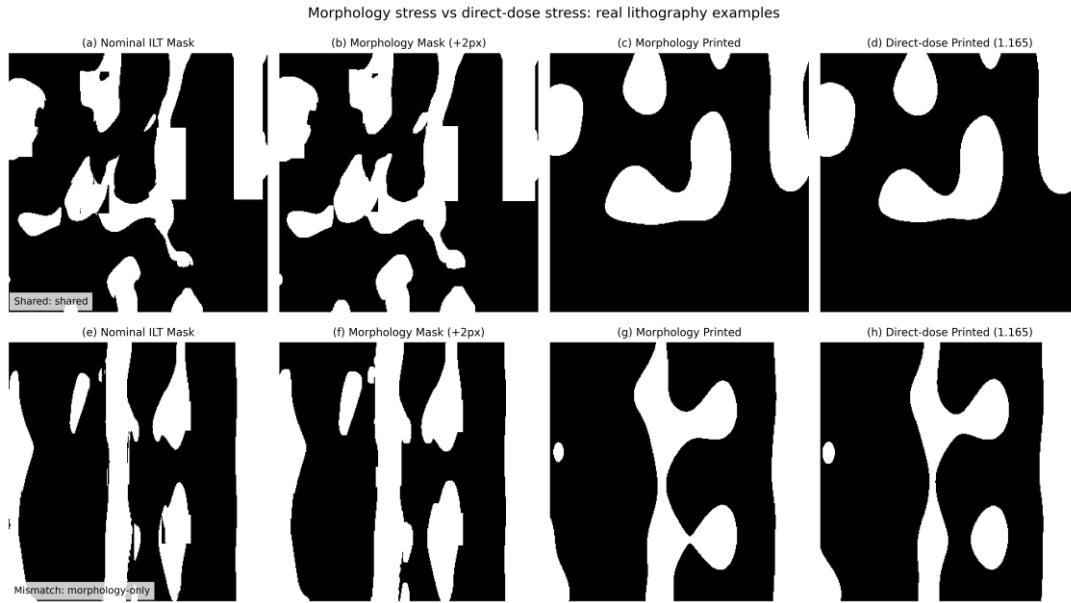


圖 17. Morphology stress 與 direct-dose stress 的 shared failure 與 mismatch 代表案例

事件匹配的位置容差由 32 奈米掃描至 512 奈米時，Jaccard 僅由 0.19 上升至 0.30，顯示兩種 stress mode 的差異不能由定位容差單獨解釋。後續工作可直接建模多個 process condition，以取代將所有變異壓縮為單一 morphology 或 dose 軸的表示方式。

4.7 綜合討論

4.7.1 三個研究問題的回答

Q1 的結果支持全域成像指標與事件指標應並列。DOINN 在壓力集的全域 IoU 為 0.867，高於 DAMOLitho 的 0.799，但橋接 recall 分別為 0.501 與 0.785，模型排序不同。邊緣帶 IoU 對橋接的相關性較高，也未在斷裂與消失兩類呈現一致優勢。因此，將單一像素指標作為所有缺陷可靠度的代理仍不足。Q2 的主要證據來自相同 DOINN 下的資料對照。CtrlSet 的橋接 recall 為 0.566，MixSet 的三次平均為 0.8634；結合近臨界間隙稀少的幾何統計，支持補充缺陷分佈的重要性。LOOSet 對 DOINN 與 TopoNet 的影響不同，說明資料涵蓋程度與架構可能交互作用，不能推論模型完全不依賴擾動分佈。Q3 的答案同時包含改善與代價。TopoNet 相對同樣使用 MixSet 的 DOINN，在橋接平均 recall 與 F1 上分別增加 0.0823 與 0.0262；下表以相同 571 個參考事件呈現原始計數的平均變化，使增益可被對應到實際漏檢與誤報數量。

表 40. TopoNet 相對增強 DOINN 的橋接改善與代價

指標	DOINN+增強	TopoNet	TopoNet - 增強
平均 TP	493.00	540.00	+47.00
平均 FN	78.00	31.00	-47.00
平均 FP	8.67	31.33	+22.67
Precision	0.9829	0.9453	-0.0375
Recall	0.8634	0.9457	+0.0823
F1	0.9192	0.9454	+0.0262

註：各欄為三個隨機種子的算術平均；差值由未四捨五入的數值計算。計數平均不是單一模型的整數事件數。三次模型各自面對相同的 571 個參考事件。

相對增強 DOINN，TopoNet 的平均漏檢由 78 個降至 31 個，減少約 60.3%；同時平均誤報增加約 22.7 個。這組結果表明 TopoNet 對橋接事件更敏感，且在本測試的 F1 上取得增益。漏檢與誤報在後續流程中的成本未於本研究定義，因此不能單由 F1 或相對漏檢降幅決定實際部署策略。

4.7.2 架構增益與正常成像的取舍

僅加入小波的 WRD-Net 橋接平均 recall 為 0.848 ± 0.039 ，未呈現高於增強 DOINN 的平均結果，因此目前證據支持的是 TopoNet 組合設計的效益，尚不足以將全部增益歸於小波本身。單次消融中，移除非對稱損失會提高全域與邊緣帶 IoU，卻使橋接 recall 由 0.949 降至 0.907；這提供像素保真度與事件敏感度可能具有不同最佳化方向的具體例子。194 個未擾動 cell 的結果揭露另一個必要評估面向：TopoNet 的 IoU 為 0.8207，有誤報切片為 14 個，並產生 471 個新增元件。此觀察不能直接換算為 471 處短路，也不足以定位是哪個模組造成誤報，但已指出正常輸出的結構穩定性是後續改良的明確目標。表 17 的 MetalSet 正常 IoU 與此處來自不同測試集合，兩者方向不同，正說明模型品質需要按評估分佈解讀。

4.7.3 電路設計用途與方法的使用界線

Ibex 延伸顯示，事件評估可利用 net 資訊形成可定位的跨網路風險，且三個模型在各已測資料組維持改善方向。它為設計者提供的研究資訊是「哪些位置的成像預測可能漏掉連通性變化」，而完整功能後果仍需網表注入或電性模型。本文沒有量測晶圓良率，也未完成神經模型參與 ILT 迭代後的端到端品質及時間比較。依目前結果，可將缺陷敏感度與正常保真度視為需共同最佳化的目標，

並以神經預測提出候選位置、再由物理模擬複核作為後續研究方向。本文尚未驗證完整混合流程，因此不給出固定通過閾值或部署保證。其已完成的成果是可追蹤事件的評估方法、可量測的改善，以及可直接指向下一階段的誤報與成像限制。

4.8 研究限制

本研究對 DOINN 完成最完整的診斷與修復驗證，包含多個 random seed、資料量對照與擾動泛化測試。本文所觀察到的「8 倍下採樣削弱細縫」屬於本研究 DOINN 實作的具體機制，可能也出現在具有相似前端下採樣設計的模型，但不能直接推廣至所有 Fourier neural operator；GAN 類模型亦可能具有不同的失效機制。其餘三個模擬器未納入與 DOINN 同等完整的 retraining/ablation，主要受單 GPU 計算資源限制：CFNOLitho 的複數運算使單一 epoch 約需 10.8 小時，DAMOLitho 與 LithoGAN 的對抗訓練成本更高。因此表 11 主要使用官方預訓練權重作為跨架構參照，不能將 DOINN 上的根因診斷直接視為三者已被逐一驗證。物理模型方面，本研究採用 LithoBench 提供的簡化模型：24 個光學 kernel、binary mask 與簡化光阻轉換，focus 亦僅有有限離散狀態。數值孔徑、照明等完整校正參數未在目前資料中提供，因此本研究不將結果宣稱為經 foundry 校準的 45 nm lithography process；Nangate45 在本文代表版圖／標準元件 design platform。所有神經模型使用相同物理 reference、相同 mask 與相同版圖，可降低物理模型差異造成的混淆，但不能完全消除 simulator fidelity 對絕對失效數量的影響。光罩生成方面，gradient ILT 未包含完整 MRC。topology-constrained 版本的印出 IoU 中位數為 0.660，原始版本為 0.784 且有 8 個 nominal artifact，顯示拓樸保護與像素保真度的取捨。NeuralILT 跨域的 85.7% 平均 target coverage 未涵蓋額外印刷、net-level topology 或 MRC，不能用此單一指標判定方法優劣。各模型在同一評估組使用相同光罩，但不同版本的事件數與 recall 不直接混合。Ibex 的 net 歸屬來自 DEF；標準元件則依 LEF 可識別 pin 區域分析，未涵蓋完整內部連線或 LVS。布林抽象未建模橋接電阻、驅動衝突與時序效應。本文的主要貢獻是預測品質分析與改善，尚未提供完整的神經／物理推論時間對照、ILT 迭代閉環驗證或晶圓製造量測。模型可作為快速候選評估的研究基礎，實際節省的物理計算及能否改善電路成功率仍屬後續驗證問題。

第五章 結論與未來發展

5.1 結論

本研究以 LithoBench 為核心，從成像品質、近臨界資料分佈及拓樸感知模型三個研究問題建立分析與改善流程，並以 Ibex 檢查電路連線層級的延伸用途。主要結論如下。

（一）成像相似度與缺陷可靠度需要分別量測。壓力集上 DOINN 與 DAMOLitho 的 IoU 和橋接 recall 排序不同；低擾動餘裕位置亦更容易漏檢。逐像素、邊緣與事件級指標共同描述模型的適用範圍。

（二）物理擾動增強能補充基準資料欠缺的近臨界案例。DOINN 的橋接 recall 由重複訓練基準 0.520 提升至 0.8634；等量重複抽樣僅取得 0.566，支持缺陷分佈比單純增加重複資料更有幫助。此結論以已測架構、分割及擾動條件為範圍。

（三）TopoNet 在共同 MixSet 下提供額外事件改善。橋接平均 recall 與 F1 分別為 0.9457、0.9454；相對增強 DOINN，每次面對相同 571 個參考事件時，平均少漏 47 個事件、增加約 22.7 個誤報。斷裂單次 recall 與 F1 為 0.8747、0.8898，補充了橋接以外的事件證據。

（四）可靠度提升具有明確的分佈與指標條件。TopoNet 在 194 個未擾動 cell 上的 IoU 為 0.8207，有誤報切片為 14 個，顯示正常成像保真度與缺陷敏感度尚有取捨。這項發現將後續目標具體化為同時改善事件品質與正常輸出穩定性。

（五）事件配對可延伸至電路連線風險。Ibex 主測試及跨層／跨設計組維持相同改善方向；DEF／LEF 資訊可區分同網路融合與跨網路短路。這建立了由成像分析走向電路判斷的研究基礎，完整功能效應、ILT 迭代效益及製造量測則保留為後續驗證。

5.2 未來發展

第一，優先改善正常樣本與壓力樣本的共同表現。可調整混合訓練比例、非對稱損失與細小元件正則化，並以正常 IoU、事件級 precision／recall／F1 與有誤報切片比例共同評估。閾值或模型選擇應在獨立驗證集完成，避免僅追求壓力

集 recall。第二，建立可輸入 dose、focus 等條件的神經微影模擬器，直接學習不同成像條件，降低對形態學代理的依賴；同時比較物理參考改變時的成像與事件品質。第三，在共同切片、物理參考與評估規則下，比較 NeuralILT、gradient ILT 與 topology-constrained ILT 的印出 IoU、EPE、net-level topology、製程餘裕及 MRC。若將神經模型放回 ILT/OPC 迭代，應量測它是否能正確追蹤中間光罩的缺陷與改善方向，而非只看最終影像相似度。第四，在資源允許時，對其他神經架構及各消融模組執行重複訓練，檢查改善能否跨架構維持；另補充推論時間與物理複核比例，評估神經/物理混合驗證的實際效率。第五，與電路設計研究結合，將可定位的跨網路事件映射至網表重要度，挑選需優先物理複核的位置；對選定事件執行 gate-level fault injection，或以 SPICE 分析橋接電阻與驅動衝突。下一階段可逐步由「成像準確」走向「電路判斷準確」，但需以新增的電路評估證據支持，不能直接將本報告的 recall 換算為下線成功率。

參考文獻

- [1] S. Zheng, H. Yang, B. Zhu, B. Yu, and M. D. F. Wong, “LithoBench: Benchmarking AI Computational Lithography for Semiconductor Manufacturing,” *Advances in Neural Information Processing Systems*, vol. 36, Datasets and Benchmarks Track, 2023, doi: 10.52202/075280-1315.
- [2] H. Yang, Z. Li, K. Sastry, S. Mukhopadhyay, M. Kilgard, A. Anandkumar, B. Khailany, V. Singh, and H. Ren, “Generic Lithography Modeling with Dual-band Optics-Inspired Neural Networks,” *Proc. 59th ACM/IEEE Design Automation Conference (DAC)*, pp. 973–978, 2022.
- [3] W. Ye, M. B. Alawieh, Y. Lin, and D. Z. Pan, “LithoGAN: End-to-End Lithography Modeling with Generative Adversarial Networks,” *Proc. 56th ACM/IEEE Design Automation Conference (DAC)*, Art. no. 107, 2019.
- [4] G. Chen, W. Chen, Q. Sun, Y. Ma, H. Yang, and B. Yu, “DAMO: Deep Agile Mask Optimization for Full-Chip Scale,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 9, pp. 3118–3131, 2022, doi: 10.1109/TCAD.2021.3116511.
- [5] B. Jiang, L. Liu, Y. Ma, H. Zhang, B. Yu, and E. F. Y. Young, “Neural-ILT: Migrating ILT to Neural Networks for Mask Printability and Complexity Co-optimization,” *Proc. IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2020, doi: 10.1145/3400302.3415704.
- [6] Z. Li et al., “Fourier Neural Operator for Parametric Partial Differential Equations,” *Proc. International Conference on Learning Representations (ICLR)*, 2021.
- [7] T. Ajayi et al., “OpenROAD: Toward a Self-Driving, Open-Source Digital Layout Implementation Tool Chain,” *Proc. Government Microcircuit Applications and Critical Technology Conference (GOMACTech)*, pp. 1105–1110, 2019.
- [8] Nangate Inc., “NanGate FreePDK45 Open Cell Library,” 2011.
- [9] A. Poonawala and P. Milanfar, “Mask Design for Optical Microlithography—An Inverse Imaging Problem,” *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 774–788, 2007, doi: 10.1109/TIP.2006.891332.
- [10] J.-R. Gao, X. Xu, B. Yu, and D. Z. Pan, “MOSAIC: Mask Optimizing Solution with Process Window Aware Inverse Correction,” *Proc. 51st ACM/EDAC/IEEE Design Automation Conference (DAC)*, 2014, doi: 10.1145/2593069.2593163.
- [11] lowRISC, “Ibex: An embedded 32-bit RISC-V CPU core,” open-source project.
- [12] S. Mallat, “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [13] H. Yang, Y. Lin, B. Yu, and E. F. Y. Young, “Lithography Hotspot Detection: From Shallow to Deep Learning,” *Proc. IEEE International System-on-Chip Conference (SOCC)*, pp. 233–238, 2017.
- [14] X. Ma and G. R. Arce, *Computational Lithography*. Hoboken, NJ, USA: John Wiley & Sons, 2010.
- [15] A. Korovin, I. Fedorov, A. Vasilev, I. Gerasimov, S. Sedunov, and S. Budenny, “Topology-inspired metric for detecting potential defects in lithography,” *Expert Systems with Applications*, vol. 323, Art. no. 132161, 2026, doi: 10.1016/j.eswa.2026.132161.
- [16] S. Zheng, B. Yu, and M. D. F. Wong, “OpenILT: An Open Source Inverse Lithography Technique Framework,” *Proc. 2023 IEEE 15th International Conference on ASIC (ASICON)*, 2023, doi: 10.1109/ASICON58565.2023.10396314.
- [17] S. Sun, F. Yang, B. Yu, L. Shang, and X. Zeng, “Efficient ILT via Multi-level Lithography Simulation,” *Proc. 60th ACM/IEEE Design Automation Conference (DAC)*, pp. 1–6, 2023, doi: 10.1109/DAC56929.2023.10247704.
- [18] G. Chen, Z. Yu, H. Liu, Y. Ma, and B. Yu, “DevelSet: Deep Neural Level Set for Instant Mask Optimization,” *Proc. IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp. 1–9, 2021, doi: 10.1109/ICCAD51958.2021.9643464.

- [19] H. Yang, S. Li, Z. Deng, Y. Ma, B. Yu, and E. F. Y. Young, “GAN-OPC: Mask Optimization With Lithography-Guided Generative Adversarial Nets,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2020, doi: 10.1109/TCAD.2019.2939329.
- [20] H. Yang and H. Ren, “Enabling Scalable AI Computational Lithography with Physics-Inspired Models,” *Proc. 28th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 715–720, 2023, doi: 10.1145/3566097.3568361.
- [21] W. Zhao et al., “AdaOPC: A Self-Adaptive Mask Optimization Framework for Real Design Patterns,” *Proc. IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2022, doi: 10.1145/3508352.3549468.
- [22] V. Singh et al., “Making a Trillion Pixels Dance,” *Proc. SPIE Optical Microlithography XXI*, vol. 6924, 2008.
- [23] Z. Yu, G. Chen, Y. Ma, and B. Yu, “A GPU-enabled Level Set Method for Mask Optimization,” *Proc. Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 1835–1838, 2021, doi: 10.23919/DATE51398.2021.9474212.
- [24] H. Yang, W. Zhong, Y. Ma, H. Geng, R. Chen, W. Chen, and B. Yu, “VLSI Mask Optimization: From Shallow to Deep Learning,” *Integration, the VLSI Journal*, vol. 77, pp. 96–103, 2021, doi: 10.1016/j.vlsi.2020.11.001.
- [25] H.-C. Shao, C.-W. Lin, and S.-Y. Fang, “Data-Driven Approaches for Process Simulation and Optical Proximity Correction,” *Proc. 28th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 721–726, 2023, doi: 10.1145/3566097.3568362.
- [26] X.-Y. Ma and S. Hao, “Inverse Lithography Physics-Informed Deep Neural Level Set for Mask Optimization,” *Applied Optics*, vol. 62, no. 33, pp. 8769–8779, 2023, doi: 10.1364/AO.503332.
- [27] Y.-H. Son, D.-H. Shin, D.-J. Lee, H. J. Lee, and T.-E. Kam, “Optical Diffraction-based Convolution for Semiconductor Lithography,” *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12458–12468, 2026.
- [28] H. He, Z. Wang, J. Wang, T. Wu, X. He, B. Yu, J. Yu, and H. Geng, “LithoSim: A Large, Holistic Lithography Simulation Benchmark for AI-Driven Semiconductor Manufacturing,” *Advances in Neural Information Processing Systems*, vol. 38, 2025.
- [29] A. B. Kahng, L. Wang, and B. Xu, “TritonRoute: The Open-Source Detailed Router,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 40, no. 3, pp. 547–559, 2021, doi: 10.1109/TCAD.2020.3003234.
- [30] A. B. Kahng, L. Wang, and B. Xu, “The Tao of PAO: Anatomy of a Pin Access Oracle for Detailed Routing,” *Proc. 57th ACM/IEEE Design Automation Conference (DAC)*, pp. 1–6, 2020.
- [31] C. Wolf and J. Glaser, “Yosys – A Free Verilog Synthesis Suite,” *Proc. Austrochip Workshop on Microelectronics*, pp. 47–52, 2013.
- [32] C. A. Mack, *Fundamental Principles of Optical Lithography: The Science of Microfabrication*. Chichester, U.K.: John Wiley & Sons, 2007, doi: 10.1002/9780470723876.
- [33] H. J. Levinson, *Principles of Lithography*, 3rd ed. Bellingham, WA, USA: SPIE Press, 2010.
- [34] A. K.-K. Wong, *Resolution Enhancement Techniques in Optical Lithography*. Bellingham, WA, USA: SPIE Press, 2001.
- [35] H. H. Hopkins, “On the Diffraction Theory of Optical Images,” *Proceedings of the Royal Society of London. Series A*, vol. 217, no. 1130, pp. 408–432, 1953, doi: 10.1098/rspa.1953.0071.
- [36] A. Mosinska, P. Márquez-Neila, M. Koziński, and P. Fua, “Beyond the Pixel-Wise Loss for Topology-Aware Delineation,” *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3136–3145, 2018.
- [37] X. Hu, F. Li, D. Samaras, and C. Chen, “Topology-Preserving Deep Image Segmentation,” *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [38] J. R. Clough, N. Byrne, I. Oksuz, V. A. Zimmer, J. A. Schnabel, and A. P. King, “A Topological Loss Function for Deep-Learning Based Image Segmentation Using Persistent Homology,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 8766–8778, 2022, doi: 10.1109/TPAMI.2020.3013679.

- [39] S. Shit et al., “clDice – A Novel Topology-Preserving Loss Function for Tubular Structure Segmentation,” Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 16560–16569, 2021, doi: 10.1109/CVPR46437.2021.01629.
- [40] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234–241, 2015.