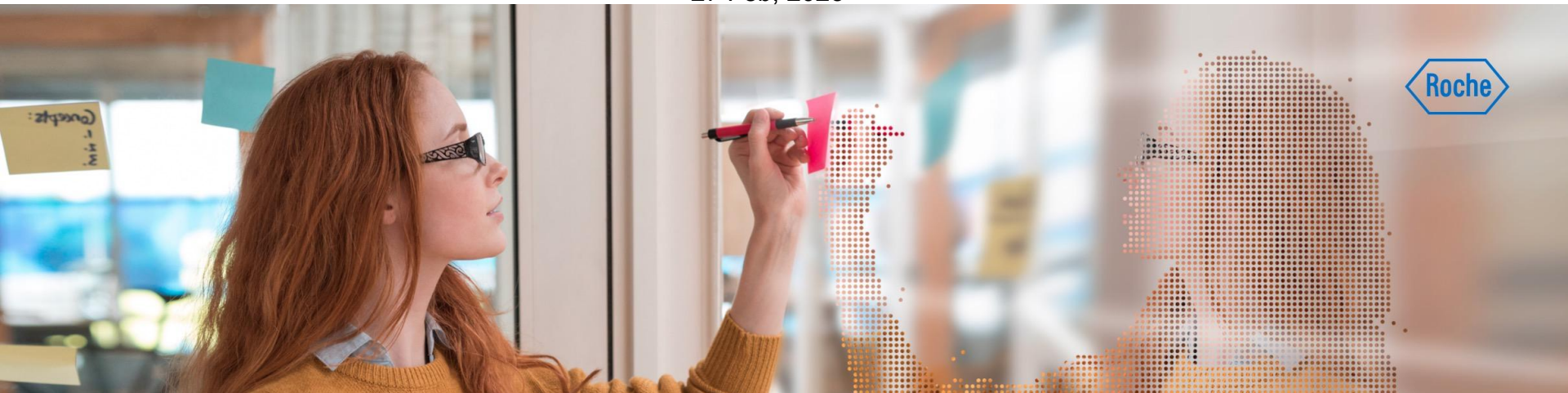
Overview of Data Science Capabilities of Roche Pharma Personalized Healthcare (PHC)

Jian Dai PhD, Yajing Zhu PhD
On behalf of Data, Analytics, and Imaging group
27 Feb, 2023



Agenda

1. Data Science in Roche Pharma PHC
2. Capabilities showcase
3. The team

Data Science in Roche Pharma PHC

Roche has consistently topped the list of pharma R&D budgets

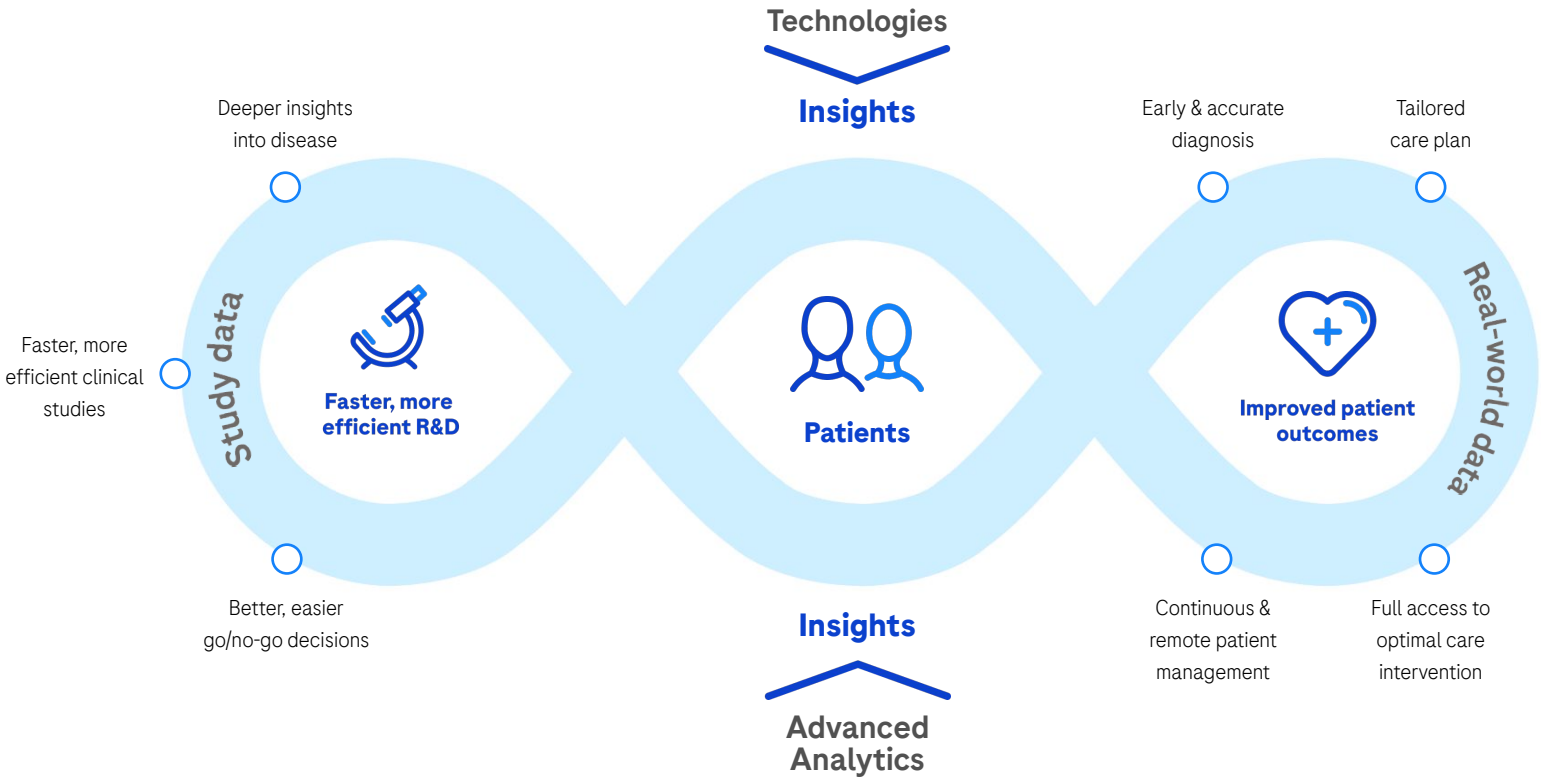
In 2021, Roche invested 14.8 B Swiss Francs ($16.1B), a 14% increase over 2020 or 23% of its revenue.

Top 10 pharma R&D budgets in 2021

Company	R&D budget ($)	Chg from 2020	2021 revenue ($)	%R&D/revenue
Roche	**16.1B**	14%	71.4B	23%
J&J	14.7B	21%	93.8B	16%
Pfizer	13.8B	47%	81.3B	17%
Merck & Co	12.2B	-9%	48.7B	25%
BMS	11.3B	2%	46.4B	24%
AstraZeneca	9.7B	62%	37.4B	26%
Novartis	9B	-6%	51.6B	17%
GSK	7.2B	4%	46.2B	16%
AbbVie	7.1B	8%	56.2B	13%
Eli Lilly	7B	15%	28.3B	25%

Data source: https://www.fiercebiotech.com/special-reports/top-10-pharma-rd-budgets-2021

Personalised Healthcare (PHC) is the confluence of advanced analytics and technology, with patients at its centre



Veeva ID: M-XX-00011617. Date of preparation: May 2022

External facing website: https://sites.google.com/roche.com/pharmaphc/about?authuser=0

Three pillars of the AI revolution

Data, Computation and Algorithms at Roche Pharma PHC

Data

- Diverse modalities of Real World Data (RWD): imaging (MRI, CT), clinical, genomics etc.
- Acquired Flatiron in 2018 for $2.1B.
- One of the founding partners of the UK INSIGHT data hub
- Multiyear partnership with Monash U to create a unique biobank (source)
- EDIS (Enhanced Data & Insights Sharing) (source)
- A global imaging platform stores clinical images of ~100k patients from 200+ studies

Computing power

- As early as 2016, Genentech research HPC cluster already had the GPU computing capabilities
- Apollo: an end-to-end platform enabling the entire ecosystem for data products development (source)
- AWS for the majority of the cloud workloads including HPC, analytics, ML, database, storage & security (source)

Algorithms

- AI solutions dev in Radiology, Ophthalmic imaging, Histopathology imaging
- Digital solutions dev in Onc/Ophtha/Neuro
- Example: 150+ manuscripts / presentations / posters since 2019 related Ophthalmology PHC
- Companion patents in close col with company IP legal

Data, Analytics, & Imaging (DAI) is the Data Science group in Roche Pharma PHC

Our Purpose : DAI brings data to life.

We enable actionable, meaningful, and novel insights that empower patients and providers to make health-related decisions that improve lives.

DAI is involved throughout the PHC product development process. We power PHC's products with core components, such as data and algorithms, that fulfil unmet needs for patients and health care providers. As experts who push the edge of our craft, we are entrepreneurial, contributing and pursuing new ideas that have the potential to make life-changing impact. Along the way, we bring our strong, connected networks, partnerships, and collaborative spirit to the work.

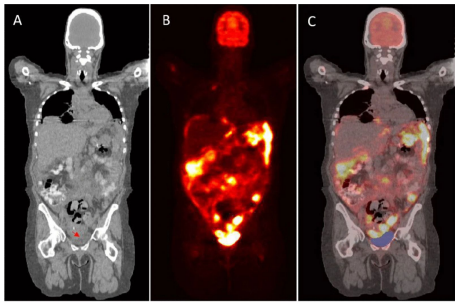
Unbridled passion and conviction around how data can improve patient lives drives us to apply our individual and collective talents and skills to the field of personalized healthcare.

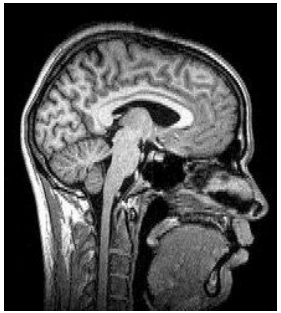
Our tech stack

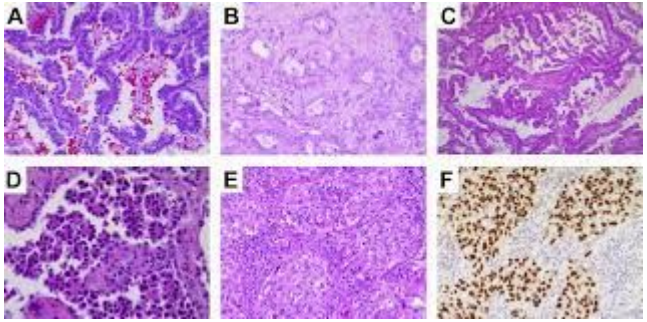
We are moving from an algo R&D team to a solution team

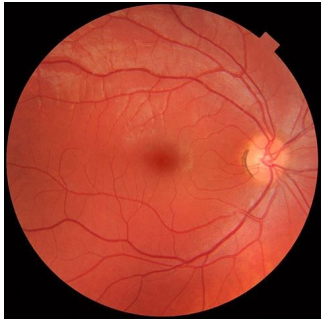
Our focused areas

AI and Digital







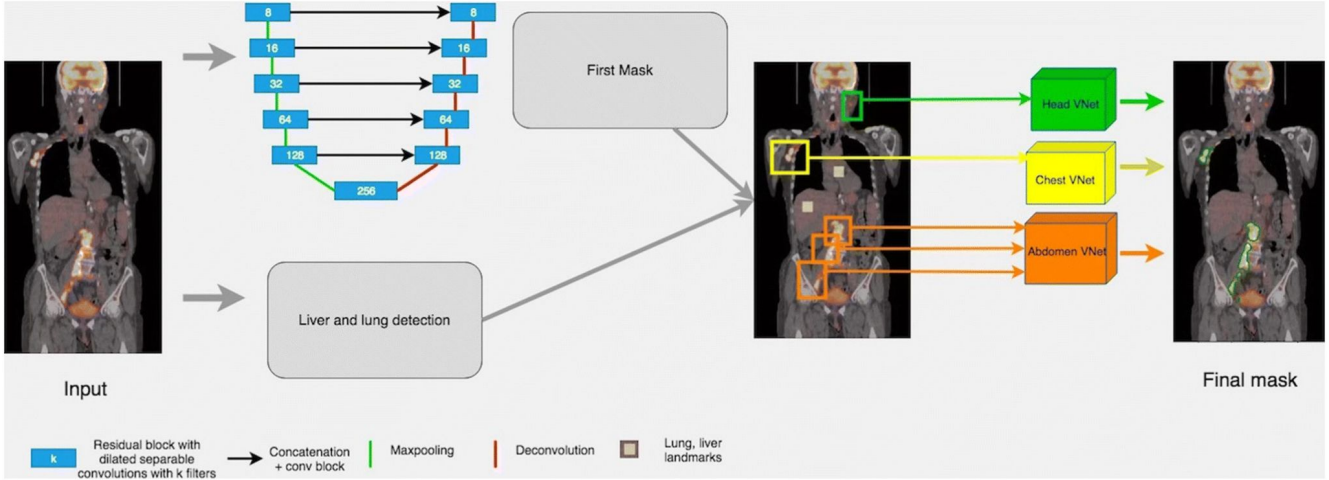


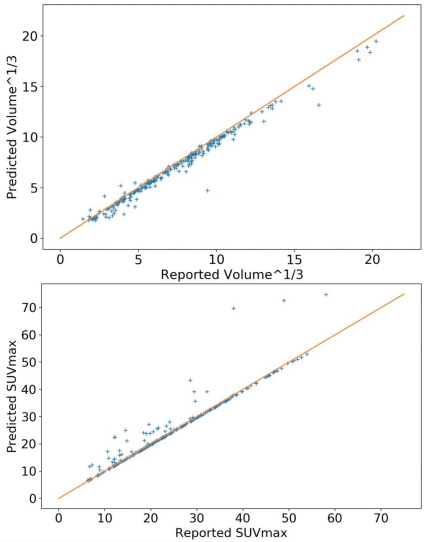
Data science capabilities showcase

Showcase: Oncology Imaging

Tumor Segmentation and Feature Extraction from Whole-Body FDG-PET/CT Using Cascaded 2D and 3D Convolutional Neural Networks

Jemaa, Fredrickson, Carano, Nielsen, Crespigny & Bengtsson, Journal of Digital Imaging vol 33, 888–894 (2020)





Dataset	Number of scans	Dice score	Sensitivity
DLBCL (training)	2266	0.895	93.2
Follicular lymphoma (test)	1124	0.886	92.6
Lung cancer (test)	274	–	93.0

Only a partial “ground truth” is available for the NSCLC test set. Thus, only sensitivity is being reported for these scans

Source: https://link.springer.com/article/10.1007/s10278-020-00341-1

Showcase: Neuroimaging

Joint MRI T1 Unenhancing and Contrast-enhancing Multiple Sclerosis Lesion Segmentation with Deep Learning in OPERA Trials

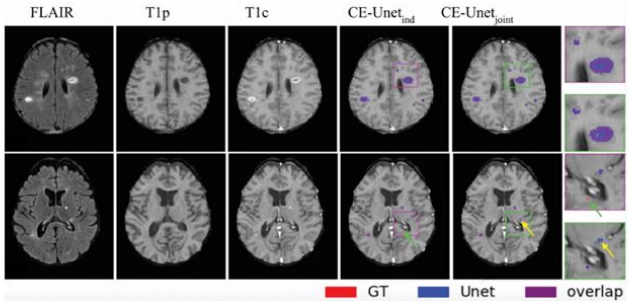
Krishnan, Song, Clayton, Gaetano, Jia, Crespigny, Bengtsson, Carano, Radiology. 2022

Call for collaboration

1. RW MRI imaging data from 1+ sites
2. Linked MRI & clinical outcomes
3. Piloting sites for algo. co-development/validation/tests

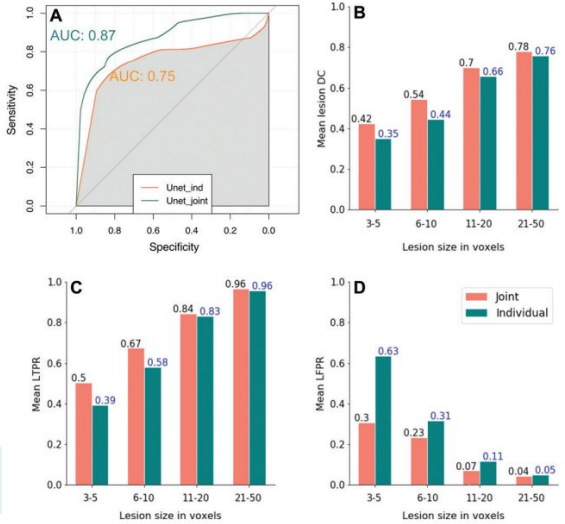
Contact: yajing.zhu@roche.com

Joint MRI T1 Unenhancing and Contrast-enhancing Multiple Sclerosis Lesion Segmentation with Deep Learning in OPERA Trials



MRI scans and overlays of ground truth (GT) and model-predicted masks show T1 unenhancing (UE) and T1 gadolinium-enhancing (CE) lesion segmentation by jointly and individually trained models. Yellow arrow = false-positive finding, green arrow = false-negative finding, FLAIR = fluid-attenuated inversion recovery, T1-c = postcontrast T1-weighted imaging, T1-p = precontrast T1-weighted imaging.

- Deep learning models developed on OPERA clinical trial MRI data sets improved segmentation and detection of enhancing multiple sclerosis lesions (true- and false-positive rates, 0.88 and 0.04, respectively).
- An ocrelizumab treatment response ($P <$ .001) reduced the mean number of T1 contrast-enhancing lesions at 24, 48, and 96 weeks (joint-model reductions: 86%, 90%, and 94%, respectively; manual assessment reductions: 93%, 96%, and 97%) and new T1 unenhancing lesions summed across all follow-up examinations (72% and 68% reductions) in the treatment arm when compared to the control arm.



Krishnan A et al. Published Online: December 14, 2021
https://doi.org/10.1148/radiol.211528

Radiology

Source: https://pubs.rsna.org/doi/10.1148/radiol.211528

Showcase: Real World Clinicogenomic Data

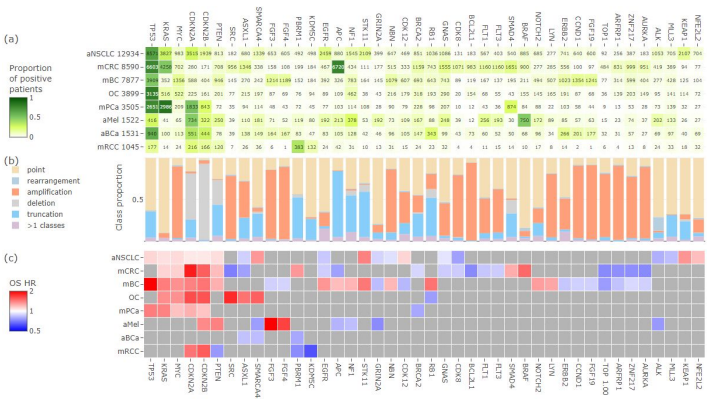
Systematic pan-cancer analysis of mutation–treatment interactions using large real-world clinicogenomics data

Liu, Rizzo, Waliany, Garmhausen, Pal, Huang, Chaudhary, Wang, Harbron, Neal, Copping, Zou, Nature Medicine 2022

Introduction

- Aim to improve outcomes & advance precision med.
- By quantifying the Tx effect of cancer therapies per tumor mutations.
- EHR data from 41K US cancer patients analyzed
- 458 mutations Id'ed predictive for survival per cancer Tx (CIT, Chemo, targeted) for 8 cancer types
- Mutation-mutation interactions characterized.

Statistics



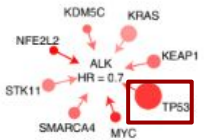
Mutation-Tx interaction

(a) aNSCLC

Immunotherapy	Chemotherapy
All (N=4072) EGFR KRAS NBN MLL3 NBN ASXL1 SRC	All (N=3446) APC EGFR STK11 BRCA2 TP53
Carboplatin, Pembrolizumab, Pemetrexed (N=1589) NFE2L2 MYC GRIN2A	Carboplatin, Paclitaxel (N=1225) GRIN2A MLL3 MYC ERBB2 CDK12
Pembrolizumab (N=1074) KRAS PBRM1 MLL3	Carboplatin, Pemetrexed (N=990) ERBB2 KRAS TP53 RB1
Nivolumab (N=529) BRAF EGFR RB1 PTEN NF1	
Carboplatin, Paclitaxel, Pembrolizumab (N=310) CDKN2B	Carboplatin, Paclitaxel Protein-Bound (N=271) KRAS TP53

Mutation-mutation interaction

a ALK - aNSCLC ($N = 1,053$)



Gene Detail: TP53

Features	Values
Anchor gene:	ALK
Cancer:	aNSCLC
Number of patients (percentage):	664 (63.1%)
HR - Anchor Positive (95% CI):	1.77 (1.32, 2.37)
Pathogenic Subtype (95% CI):	1.85 (1.38, 2.49)
Pathogenic Subtype Definition:	point, rearrangement (N=972)

Gene Detail: KRAS

Features	Values
N - given drug:	421 (39.2%)
Interaction [exp(coef)] (95% CI):	0.75 (0.62, 0.91)
HR for gene mutation (95% CI):	0.79 (0.61, 1.02)
HR for absence of gene mutation (95% CI):	0.87 (0.72, 1.05)
Mutation subtype with strongest treatment interaction:	point (97.9%)
HR for that subtype of mutations (95% CI):	0.76 (0.58, 0.99)
HR given other drug (95% CI):	0.90 (0.74, 1.11)

Source: https://www.nature.com/articles/s41591-022-01873-5 http://www.precision-cancer.org/ https://github.com/RuishanLiu/precision-cancer

The team

Our amazing team: global, fun, focused!

Wrap up

What you can take home…

1. At Roche, personalised healthcare – also known as PHC – is a key area of focus, combining pioneering science, data, analytics and technology to advance new discoveries and improve patients' lives.
2. Data, Analytics, & Imaging (DAI) is the Data Science group in Roche Pharma PHC
3. DAI focuses on the development of AI solutions in Radiology, Ophthalmic Imaging and Histopathology Imaging and Digital technologies.
4. We have demonstrable capabilities in algorithm R&D and productization.
5. We are a team full of energy, life and curiosity.
6. If you are interested in learning more about our team and our work, please feel free to contact us by dai.jian@gene.com yajing.zhu@roche.com

Doing now what patients need next

Backup slides

Showcase: Ophthalmic Imaging

Deep Learning to Predict Geographic Atrophy Area and Growth Rate from Multimodal Imaging

Anegondi, Gao, Steffen, Spaide, Sadda, Holz, Rabe, Honigberg, Newton, Cluceru, Kawcynski, Bengtsson, Ferrara, Yang, Ophthal. Retina 2022

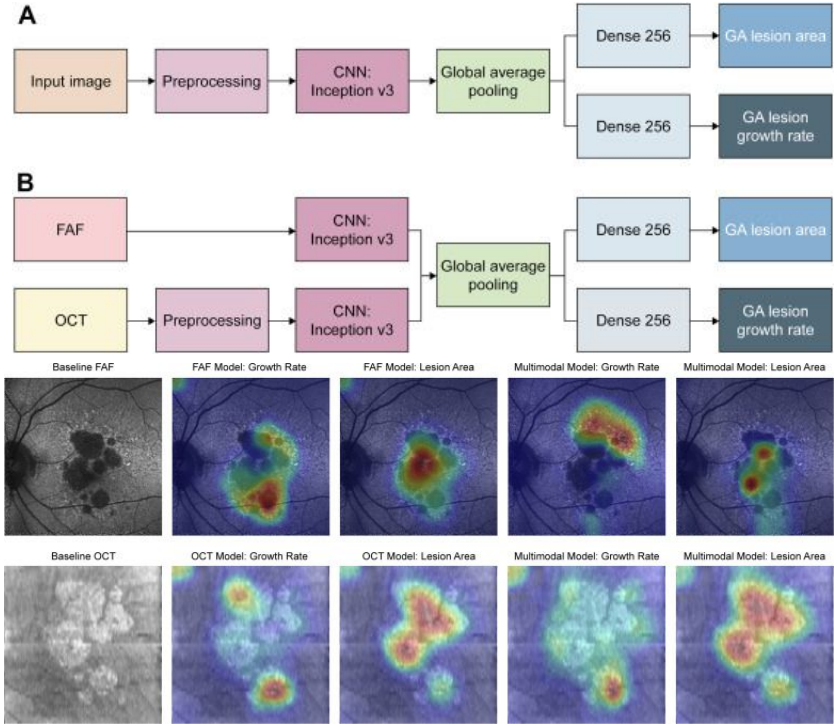


Table 2 Performance of the Benchmark Model* and 3 Multitask Models on the Development Data Set, Holdout Data Set, and Independent Test Sets

Model Type	GA Lesion Area Prediction, mm^2				GA Growth Rate, mm^2/y			
	Development Data Set (n = 1279)	Holdout Data Set (n = 443)	Independent Test Sets		Development Data Set (n = 1279)	Holdout Data Set (n = 443)	Independent Test Sets	
			MAHALO Data Set (n = 106)	Proxima B Data Set(n = 169)			MAHALO Data Set (n = 106)	Proxima B Data Set(n = 169)
Benchmark	NA	NA	NA	NA	0.25	0.16 (0.10–0.23)	0.34 (0.21–0.47)	0.25 (0.15–0.37)
FAF-only	0.93 (0.03)	0.96 (0.95–0.97)	0.98 (0.97–0.99)	0.95 (0.93–0.96)	0.48 (0.05)	0.48 (0.41–0.55)	0.65 (0.52–0.75)	0.47 (0.34–0.60)
OCT-only	0.91 (0.03)	0.91 (0.87–0.95)	NA	NA	0.42 (0.04)	0.36 (0.29–0.43)	NA	NA
Multimodal (FAF and OCT)	0.93 (0.02)	0.94 (0.92–0.96)	NA	NA	0.52 (0.05)	0.47 (0.40–0.54)	NA	NA

Source: https://www.ophthalmologyretina.org/article/S2468-6530(22)00426-2/fulltext

Showcase: Histopathology Imaging

A comparative study between vision transformers and CNNs in digital pathology

Deininger; Stimpel; Yuce; Abbasi Sureshjani; Schönenberger; Ocampo; Korski; Gaire; T4V @ CVPR 2022

1 – Introduction

- Vision transformers (ViTs) are not fully investigated in digital pathology (DP) yet compared to convolutional neural networks (CNNs)
  - No convolution-induced biases allow the ViTs to learn global features and complex relations in the data
- However, their weak inductive bias is also a limiting factor for model convergence. In DP, especially self-supervised pretraining is the most promising to foster model convergence due to the sparse availability of annotated WSIs

Emerging potential of ViTs and the need for self-supervised learning in DP

→ This work explores fully- and self-supervised (DINO) ViTs for patch-wise

- tumor detection in sentinel lymph node (SLN), diffuse large B-cell lymphoma (DLBCL), breast, and lung adenocarcinoma (LUAD) WSIs
- tissue type identification in colorectal cancer (CRC) WSIs

2 – Models

	CNN	ViT
Fully-supervised	***ResNet18*** • Pretrained on ImageNet • Trained with **Adam**	***DeiT-Tiny*** [1] • Pretrained on ImageNet • Trained with ***Sharpness-Aware Minimization* (SAM)**. • *SAM* needs two forward-backward passes for gradient descent, but strongly fostered ViT generalization.

Self-supervised	***PathNet***
	• ***ResNet18* pretrained with BYOL**
	• Pretrained on a DP dataset
	  ○ Eight tissue types: DLBCL, lymph node, follicular lymphoma, tonsil, lung, colon, breast and thyroid, for details see [2]
		DINO
		• ***DeiT-Tiny* pretrained with DINO** [3] architecture
		• Pretrained on what we call the TCGA 100 dataset
		  ○ Comprises patches from 8,747 *The Cancer Genome Atlas* (TCGA) WSIs encompassing the datasets BRCA, CHOL, HNSC, KIRC, KIRP, LIHC and PRAD.

For downstream tumor detection / tissue type identification: trained like fully-supervised models

3 – Datasets

Tissue type identification

CRC9 dataset

- 100,000 non-overlapping 224x224 color-normalized patches at magnification 20x (0.5 MPP)
- From CRC and normal tissue
- 9 classes
- 70% / 15% / 15% for training / validation / testing

Tumor detection

Extracted 256x256 tumoral and non-tumoral (normal) patches at magnification 20x (0.5 MPP)

- **SLN**: Trained / tested on 270 / 130 WSIs (Camelyon16 dataset)
- **DLBCL**: Trained / tested on 4,957 / 103 internal WSIs
- **Breast**: Trained / tested on 335 / 537 internal breast WSIs
- **LUAD**: Trained / tested on 431 / 138 internal lung WSIs

4 - Performance comparison

Very similar performance of models across all datasets

Model	FW	Metric	CRC9	SLN	DLBCL	LUAD	Breast
ResNet18	×	PR AUC	**0.999**	0.885	**0.976**	0.913	0.809
		ACC	**0.995**	0.981	**0.88**	0.858	0.915
DeiT-Tiny	×	PR AUC	0.998	**0.917**	0.97	**0.94**	0.817
		ACC	0.982	**0.988**	0.874	0.88	0.913
PathNet	×	PR AUC	**0.999**	0.908	0.97	0.92	0.818
		ACC	**0.995**	0.979	0.866	**0.885**	0.92
DINO	×	PR AUC	**0.999**	0.912	0.958	0.933	**0.828**
		ACC	0.991	0.984	0.874	0.871	**0.924**

- Also similar performance for magnification 10x, e.g. for SLN: *ResNet18* PR AUC 0.887, *DeiT-Tiny* PR AUC 0.919
- Results are on par with the state-of-the-art performances on the same DP tasks

When **fixing the weights** in the feature extraction layers, ***DINO* performed better than PathNet** for CRC9, SLN, LUAD and breast. We hypothesize this is caused by the more diverse dataset *DINO* was trained on.

Model	FW	Metric	CRC9	SLN	DLBCL	LUAD	Breast
PathNet	✓	PR AUC	0.969	0.833	**0.931**	0.887	0.711
		ACC	0.928	0.943	**0.829**	0.859	0.883
DINO	✓	PR AUC	**0.997**	**0.894**	0.887	**0.914**	**0.751**
		ACC	**0.983**	**0.959**	0.817	**0.878**	**0.897**

5 – Model comparison

Correlation of the mean test slide accuracies for SLN, DLBCL and breast with exception of a few outliers.

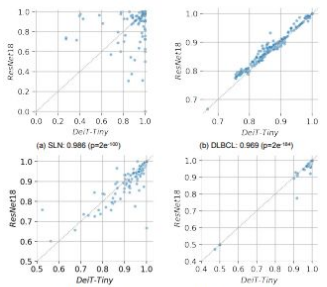


Figure 1: Mean accuracy per test slide incl. Pearson correlation.

***DeiT-Tiny* focused on more localized patch regions** in comparison to the ResNet18.

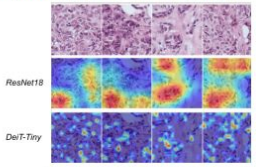


Figure 2: Comparison of *ResNet18* and *DeiT-Tiny Grad-CAM* heatmaps for randomly selected SLN tumor patches.

SAM is essential for ViT generalization, however results in **twice the runtime**

Model	#params	Throughput (img/s)
ResNet18	11M	250
DeiT-Tiny (Adam)	5M	238
DeiT-Tiny (SAM)	5M	117

6 – Conclusion

- **ViT-based models performed similar to the baseline *ResNet18*** across several datasets and across different magnifications **while requiring more training effort**
- **Correlated model predictions on slide level**
- *DINO* is more versatile than *PathNet* for tumor detection across several tissue types, however this is most likely caused by a more diverse dataset trained on
- We hypothesize that DP tissue type identification and tumor detection are tasks that can be easily learned with traditional CNN approaches
  - → ViTs often could not benefit from its larger flexibility in feature detection
  - → **We propose the application of ViTs to more challenging DP tasks**

References

[1] Hugo Touvron et al.: Training data-efficient image transformers & distillation through attention, 2020
[2] Samaneh Abbasi-Sureshjani et al.: Molecular subtype prediction for breast cancer using H&E specialized backbone, 2021
[3] Mathilde Caron et al.: Emerging properties in self-supervised vision transformers, 2021

Source: https://drive.google.com/file/d/1nXf0W-hDe3LQwa5-jQ9N-mSoPoVCFBy0/view?usp=sharing https://arxiv.org/abs/2206.00389

Showcase: ML Foundation

Unsupervised Domain Adaptation with Contrastive Learning for OCT Segmentation

Gomariz, Lu, Li, Albrecht, Maunz, Benmansour, Valcarcel, Luu, Ferrara, Goksel, MICCAI 2022

- A novel semi-supervised learning framework for segmentation of volumetric images from new unlabeled domains was proposed.
- We jointly use supervised and contrastive learning, also introducing a contrastive pairing scheme that leverages similarity between nearby slices in 3D.
- We evaluate our methods for domain adaptation from a (labeled) source domain to an (unlabeled) target domain.
- In the target domain, our method achieves a new state-of-the-art Dice coefficient and leads to results comparable to an upper bound
- In the source domain, our model also improves the results by 5.4% Dice.

Uncertainty-Aware Geographic Atrophy Progression Prediction from Fundus Autofluorescence

Yang, Anegondi, Steffen, Gao, Cluceru, Rabe, Dai, Ferrara, MACCAI 2022

- Geographic atrophy (GA) progression over time was predicted accurately using DL models with baseline fundus autofluorescence (FAF) images as input in our previous work.
- Here, both non-parametric and parametric deep ensemble approaches were applied and evaluated for the prediction uncertainty estimation using both simulated and clinical study data in a multitask regression setting.
- The results not only show promising performance in detecting near and far out-of-distribution data cases, but may also suggest the improved performance in predicting GA growth rate for in-distribution data.

Source: https://link.springer.com/chapter/10.1007/978-3-031-16452-1_34 https://link.springer.com/chapter/10.1007/978-3-031-17721-7_4