

Realizing Artificial Intelligence in Pathology

Streamlining Machine Learning Model Development and Clinical Deployment

Markus D. Herrmann, MD, PhD

June 10, 2021

Markus D. Herrmann, MD, PhD

- Interdisciplinary physician-scientist at the intersection of medicine, biology, and computer science with a specialization in biomedical imaging, machine learning, and clinical informatics.
- Studied medicine at Ulm University in Germany, where he also received his first PhD.
- Joined the faculty of MGH Pathology to lead the Computational Pathology division in 2019.
- Research focuses on the development and evaluation of digital imaging technologies and image analysis software as well as their translation into safe and effective clinical services.



Disclaimer

 The CAP does not permit reproduction of any substantial portion of the material in this Webinar without its written authorization. The CAP hereby authorizes attendees of the CAP Webinar to use the PDF presentation solely for educational purposes within their own institutions. The CAP prohibits use of the material in the Webinar – and any unauthorized use of the CAP's name or logo – in connection with promotional efforts by marketers of laboratory equipment, reagents, materials, or services.

Disclaimer

 Opinions expressed by the speaker are the speaker's own and do not necessarily reflect an endorsement by the CAP of any organizations, equipment, reagents, materials, or services used by participating laboratories.

Disclaimer

• I have no conflicts of interests related to this presentation.



- 1. Limitations of current approach to machine learning model development and deployment
- 2. Proposed solution to develop turnkey models that can be readily deployed for clinical application
- 3. Proof-of-concept use case



Ingredients of a machine learning project



The ideal machine learning project



The computational pathology gap



Data communication in machine learning



Bridging the data communication gap

Goal

Develop turnkey, deployment-ready models (plug & play)





Proposed solution

- Apply consistent data management throughout the model development and deployment process
- Rely on a standard interface to achieve interoperability

"The ability of two or more systems or components to exchange information and to use the information that has been exchanged."

Institute of Electrical and Electronics Engineers (IEEE) Standard Computer Dictionary

"Secure exchange of electronic health information with, and use of electronic health information from, other health information technology without special effort on the part of the user."

21st Century Cures Act, Section 4003 (A)



- 1. Limitations of current approach to machine learning model development and deployment
- 2. Proposed solution to develop turnkey models that can be readily deployed for clinical application
- 3. Proof-of-concept use case



Digital Imaging and Communications in Medicine



Herrmann et al. 2018

Image annotations and model outputs in DICOM



© College of American Pathologists

DICOM-enabled machine learning platform



© College of American Pathologists

10 June 2021 17

Demo of slim viewer with Google Healthcare API



© College of American Pathologists

The future of bioimaging research in the cloud

Image Data Repository

- Radiological images
- Slide microscopy images
- Image annotations
- Image-derived measurements

Image Management

DICOMweb interface of the Google Healthcare API



https://datacommons.cancer.gov



- 1. Limitations of current approach to machine learning model development and deployment
- 2. Proposed solution to develop turnkey models that can be readily deployed for clinical application
- 3. Proof-of-concept use case



Lung cancer detection in whole slide images



Collections of whole slide images in DICOM format

\rightarrow C	ŵ	🛛 🖨 https://portal.imaging.data	commons.cancer.gov/collections/			··· ⊘ ☆ ⊻	III\ 🗉 🛎 😴 🗭 🔳		
H N	ATIONAL CANCER INST naging Data Comi	nons Collections	Exploration Discourse C	Documentation 🖓			•		
•	TCGA-LUSC	Original	Lung Squamous Cell Carcinoma	CT, NM, PT, Pathology	37	Lung	Clinical, Genomics		
	DOI: 10.7937/K9/TCIA.2016.TYGKKFMQ The Cancer Genome Atlas-Lung Squamous Cell Carcinoma (TCGA-LUSC) data collection is part of a larger effort to enhance the TCGA http://cancergenome.nih.gov/ data set with characterized radiological images. The Cancer Imaging Program (CIP) with the cooperation of several of the TCGA tissue-contributing institutions are working to archive a large portion of the radiological images of the LUSC cases. Please see the TCGA-LUSC wiki page to learn more about the images and to obtain any supporting metadata for this collection.								
D	TCGA-KIRP	Original	Kidney Renal Papillary Cell Carcino	ma CT, MR, PT, Pathology	33	Renal	Clinical, Genomics		
>	TCGA-THCA	Original	Thyroid Cancer	CT, PT, Pathology	6	Thyroid	Clinical, Genomics		
•	TCGA-SARC	Original	Sarcomas	CT, MR, Pathology	5	Chest-Abdomen-Pelvis, Leg, TSpine	Clinical, Genomics		
	TCGA-ESCA	Original	Esophageal Carcinoma	CT, Pathology	16	Esophagus	Clinical, Genomics		
)	TCGA-CESC	Original	Cervical Squamous Cell Carcinoma Endocervical Adenocarcinoma	and MR, Pathology	54	Cervix	Clinical, Genomics		
)	TCGA-STAD	Original	Stomach Adenocarcinoma	CT, Pathology	46	Stomach	Clinical, Genomics		
	TCGA-COAD	Original	Colon Adenocarcinoma	CT, Pathology	25	Colon	Clinical, Genomics		
	TCGA-KICH	Original	Kidney Chromophobe	CT, MR, Pathology	15	Kidney	Clinical, Genomics		
	TCGA-READ	Original	Rectum Adenocarcinoma	CT, MR, Pathology	3	Rectum	Clinical, Genomics		
	ISPY1 (ACRIN 6657)	Original	Breast Cancer	MR, SEG	222	Breast	Clinical, Image Analyses		
	QIN-HeadNeck	Original	Head and Neck Carcinomas	PT, CT, SR, SEG, RWV	156	Head-Neck	Clinical, Image Analyses		
•	LIDC-IDRI	Original	Lung Cancer	CT, CR, DX	1010	Chest	Image Analyses, Clinical		
•	TCGA-LUAD	Original	Lung Adenocarcinoma	CT, PT, NM, Pathology	69	Chest	Clinical, Genomics, Image Analyses		
	DOI: 10.7937/K9/TCIA Note: This collection	DOI: 10.7937/K9/TCIA.2016.JGNIHEP5 Note: This collection has special restrictions on its usage. See Data Usage Policies and Restrictions.							

tias Data Portal to research the link between clinical phenome and tissue genome.

https://portal.imaging.datacommons.cancer.gov

Collections from The Cancer Imaging Archive

TCGA (n=3114)

- LUAD (n=1541)
- LSCC (n=1573)

CPTAC (n=790)

- LUAD (n=496)
- LSCC (n=294)



DICOM-based model training and inference



© College of American Pathologists

Model outputs for selected tumor tissue sections



Tissue





Tissue



35





35

Model outputs for selected tumor tissue section



Tissue





4.16 %

LUAD



58.41 %

NSCLC







54.25 %







26

10 June 2021

© College of American Pathologists

Exchange of model outputs in DICOM format



Region of Interests (ROI) Frame of reference

Qualitative Evaluations Coded concepts



https://whobluebooks.iarc.fr/

Measurements Units

Encoding of model outputs in DICOM

```
import numpy as np
import highdicom as hd
from pydicom.sr.codedict import codes
```

```
# Encode bounding box enclosing the tumor region using 3D spatial coordinates
region_item = hd.sr.ImageRegion3D(
   graphic_type=hd.sr.GraphicTypeValues3D.POLYGON,
   graphic_data=np.array([
        [45., 34., 0.],
        [49.. 34.. 0.].
        [49., 42., 0.],
       [45., 42., 0.],
        [45., 34., 0.]
   1).
   frame of reference uid="1.2.3.4"
# Encode a qualitative evaluation describing the "morphology" as "adenocarcinoma"
morphology_item = hd.sr.CodeContentItem(
   name=codes.SCT.Morphology,
   value=hd.sr.CodedConcept(
       value="8140/3".
       scheme designator="ICD03".
       meaning="Adenocarcinoma, NOS"
   )
# Encode a measurement with name "area" and of unit "square millimeter"
area item = hd.sr.NumContentItem(
   name=codes.SCT.Area,
   value=12.07,
   unit=codes.UCUM.SquareMillimeter
```

import highdicom as hd from pydicom.dataset import Dataset from pydicom.sr.codedict import codes

```
def load_image() -> Dataset: ...
```

```
source_image = load_image()
```

Describe the device that is reporting the observation observation_context = hd.sr.ObservationContext(observer_device_context=hd.sr.ObserverContext(observer_type=codes.DCM.Device, observer_identifying_attributes=hd.sr.DeviceObserverIdentifyingAttributes(uid=hd.UID(), name="GPU01"))

```
# Describe the region of interest finding and associated measurements and evaluations
roi_description = hd.sr.PlanarROIMeasurementsAndQualitativeEvaluations(
    finding_type=codes.SCT.Neoplasm,
```

```
finding_sites=[hd.sr.FindingSite(anatomic_location=codes.SCT.Lung)],
```

```
referenced_region=region_item,
measurements=[area_item],
qualitative_evaluations=[morphology_item]
```

```
# Construct a measurement report from the ROI(s)
measurement_report = hd.sr.MeasurementReport(
    observation_context=observation_context,
    procedure_reported=codes.SCT.ImagingProcedure,
    imaging_measurements=[roi_description]
```

```
# Construct an SR document containing the measurement report
report_document = hd.sr.Comprehensive3DSR(
    evidence=[source_image],
    content=measurement_report[0],
    series_number=1,
    series_instance_uid=hd.UID(),
    sop_instance_uid=hd.UID(),
    instance_number=1,
    manufacturer="MGH Computational Pathology",
    manufacturer="MGH Computational Pathology",
    software_version="V1",
    device_serial_number="XYZ"
```

Review of model outputs in slim viewer



© College of American Pathologists

DICOM-enabled machine learning platform



© College of American Pathologists

10 June 2021 30

Open-source software available on GitHub

Image: Model Computation Department of Pathology, Mass ⊙ Boston, MA Image: Repositories 10 Packages A P	reople 5 A Teams 2 III Projects 🕸	Settings
Pinned repositories		Customize pinned repositories
Image: Slim Image: Slim Interoperable slide microscopy viewer for thin clients TypeScript ☆ 7 ♀ 4	□ dicom-microscopy-viewer :: Web-based viewer for DICOM Visible Light Whole Slide Microscopy Images ● JavaScript ☆ 46 ♀ 13	□ dicomweb-client :: Python client for DICOMweb RESTful services ● Python ☆ 66 ♀ 20
☐ highdicom II High-level DICOM abstractions for the Python programming language Python ● Python ☆ 39 ♀ 8		

Participation in DICOM WG-26 Connectathon



© College of American Pathologists

DICOM WG-26 Pathology

Digital Imag	Digital Imaging and Communications in Medicine Q. Search Login									
MENU	ABOUT	STANDARD	ΑCTIVITY	USING DICOM	RESOURCES	CONFERENCES	CALENDAR	NEWS	CONTACT	
DICOM /	Activity / V	Vorking Groups & Mir	nutes / WG-26: Pa	thology						
WG-	26: Pa	thology								
Secret	ariat (U.S)		Co	ollege of American evin Schap, wg26cl	Pathologists (CA hairs@dicomsta	P) andard.org				
Secret	ariat (Europ	pe)	Sc	ociedad Española d arcial García Rojo,	e Informática de wg26chairs@di	e la Salud (SEIS) comstandard.org				
Co-Cha	airs		M M	arkus Herrmann, M ikael Wintell, Västr g26chairs@dicom	Aassachusetts G a Götalands Reg standard.org	eneral Hospital ionen				
Last st	rategy upd	ate	20	19-10-02						
Minute	25		w	G-26 minutes						
Scope	:									
• WG2 com thos	6's scope is es to produ e system w	s to support ar uce, store and o hen it comes t	d develop the communicate. o interoperabi	DICOM Standard s We also want to m lity using global Sta	so the Pathology ake it possible fo andards to creat	Domain can handle or the domain to be e workflows.	e their whole sl e a part of the w	ide image: /hole healt	s and also the macr thcare process by o	os when it pen up

DICOM WG-26 Hackathon

Hello, world!

Pathology Hackathon



Develop **interoperable computational tools** for whole slide image annotation and analysis

May 1st - June 18th 2021 (finale at ECDP 2021 congress)

https://github.com/DICOMWG26AnnotationAHG/hackathon

Collaborative development of open standards

Code () Issues 4	ាំ Pull requests 🛛 🖓 Discussions	🕞 Actions 🛄 Projects 🛄 Wiki 🕕 Security 🗠	Insights 🔅 Settin			
	양 main → 양 1 branch 🛇 0 tags	Go to file Add f	ile ▼			
	+ hackermd Add documentation of data	collections and origin servers (#7) deda537 13 days	s ago 🛛 11 commits			
	.github/ISSUE_TEMPLATE	FEMPLATE Update README and issue template				
	README.md	Add documentation of data collections and origin servers (#7)				
	🗋 data.md	Add documentation of data collections and origin servers (#7)	rs (#7) 13 days ago			
	🗅 servers.md	Add documentation of data collections and origin servers (#7)	on of data collections and origin servers (#7) 13 days ag			
	E README.md		Ô			
	DICOM WG-26 W	SI Annotation Hackathon				
	A virtual Hackathon for collaborative whole slide image (WSI) annotations and individuals) can participate free	e development and testing of software for the generation, exch in DICOM SR, SEG or Sup 222 format, in which anyone (vend ly and openly.	hange, and use of ors, academics			
	The event will be held virtually over	the internet and coordinated via this repository. It will start Ma	y 1st 2021 and			

Integrating the Healthcare Enterprise (IHE)

IHE PaLM

Digital Pathology Image Acquisition



Dash et al. 2021

Acknowledgements

MGH Pathology

- Christopher Gorman
- o Aidan Stein
- Veronica Klepeis
- o Jochen Lennerz
- Lida Harriri
- Long Le
- o John lafrate
- o David Louis

MGH Radiology / MGB CCDS

- Christopher Bridge
- Sean Doyle
- Jayashree Kalpathy-Cramer

BWH Radiology

- Andrey Fedorov
- Ron Kikinis
- Isomics
 - o Steven Pieper
- Radical Imaging
 - Igor Octaviano
 - o Davide Punzo
 - Eric Ziegler
- Google
 - o Jason Klotzer
 - Craig Mermel

DICOM WG-26

- David Clunie (Pixelmed)
- Gábor Garányi (3DHistech)
- Steve Moore (WashU)
- Norman Zerbe (Charité)
- Joachim Schmidt (Roche)
- Kevin Schap (CAP)
- IHE PaLM
 - Rash Dash (Duke)
 - Mary Kennedy (CAP)
- National Cancer Institute
 - o Ulrike Wagner
 - Keyvan Farahani

Resources of Digital and Computational Pathology Committee

- List resources
 - SPECs
 - Resources Guides
 - Topic Center Pages & Al pages
 - <u>Al@CAP.ORG</u> email address



THANK YOU!

Thank you for attending our webinar "Realizing Artificial Intelligence in Pathology – Streamlining Machine Learning Model Development and Clinical Deployment" by Markus D. Herrmann, M.D., Ph.D.

For comments about this webinar or suggestions for upcoming webinars, contact <u>Al@cap.org</u>

NOTE: There is no CME/CE credit available for today's complimentary webinar. The pdf of the presentation will be sent out in about 1 week.

