Memorial Sloan Kettering Cancer Center

Artificial Intelligence and Sarcoma Diagnosis

By Konstantinos Linos MD, FCAP, FASDP
Bone, Soft Tissue and Dermatopathology
Associate Attending
Memorial Sloan Kettering Cancer Center
Department of Pathology and Laboratory Medicine
NY, USA

- I have no relevant financial disclosures

Artificial intelligence and computational pathology

Miao Cui[1] · David Y. Zhang[2]

Laboratory Investigation (2021) 101:412–422

- **Artificial Intelligence**: A branch of computer science dealing with tasks that normally require human intelligence

- **Machine learning**: A branch of AI in which statistical algorithms establish their own patterns by being exposed to representative data to interpret and act on new data

- **Deep neural networks:** Also called deep learning, which is a subset of machine learning using complex multilayered architectures including multiple hidden layers and a large number of nodal connections

- **Artificial neural networks:** A set of layered, interconnected artificial neurons based on deep neural networks to explore higher levels features, mimicking biologic brain
  - The number of trainable parameters >100.000

- **Convolutional neural networks**: A type of artificial neural network particularly designed for machine vision field. They have been most commonly applied to analyze images such as image recognition and classification

- **GoogLeNet**: A convolutional neural network that was created by Google for computer vision and classification
- **FaceNet**: A convolutional neural network for face recognition and classification

- **Area under receiver operating curve (AUC)**: Performance measured by the area under the receiver operating characteristic curve (from 0.5 (lowest) to 1.0 (highest))

"To minimize the mistakes your AI will make, you should use the most accurate Machine Learning model."

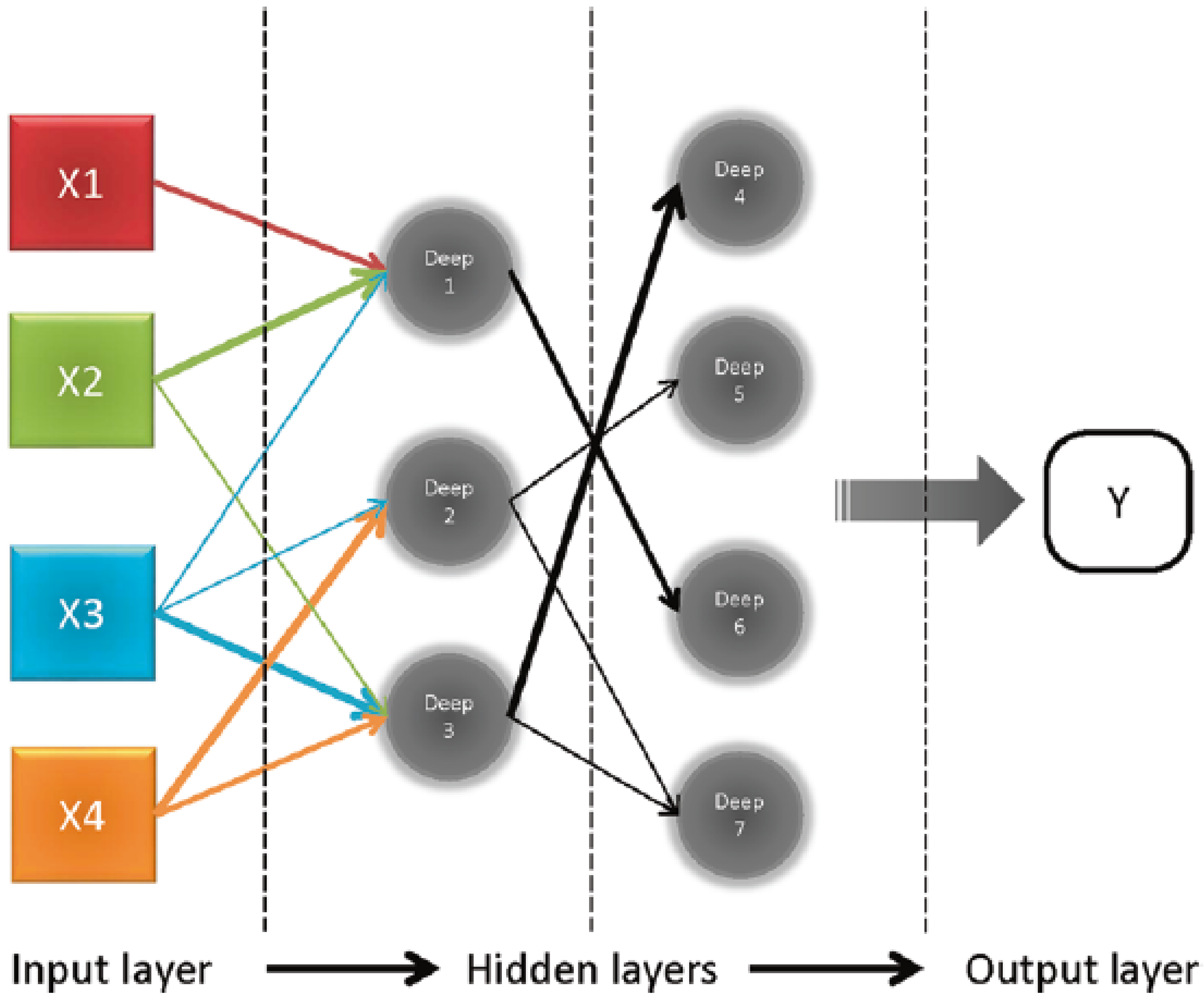
Sounds straightforward, right?

However, making the least mistakes should not always be your goal since **different *types* of mistakes can have varying impacts**. ML models will make mistakes and it is, therefore, crucial to ***decide which mistakes you can better live with***.

To choose the right ML model and make informed decisions based on its predictions, it is important to understand different measures of relevance.

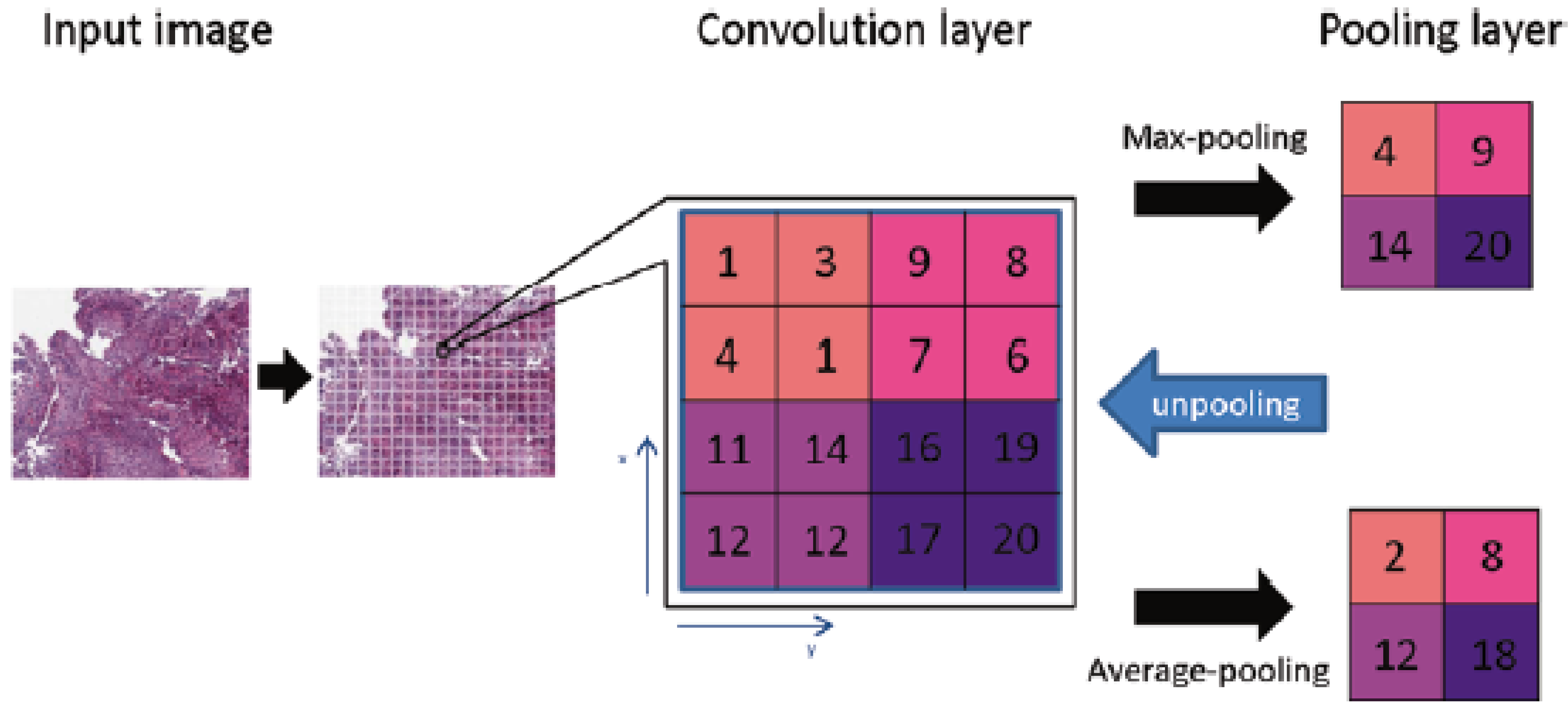
https://levity.ai/blog/precision-vs-recall

Basic structure of a deep neural network



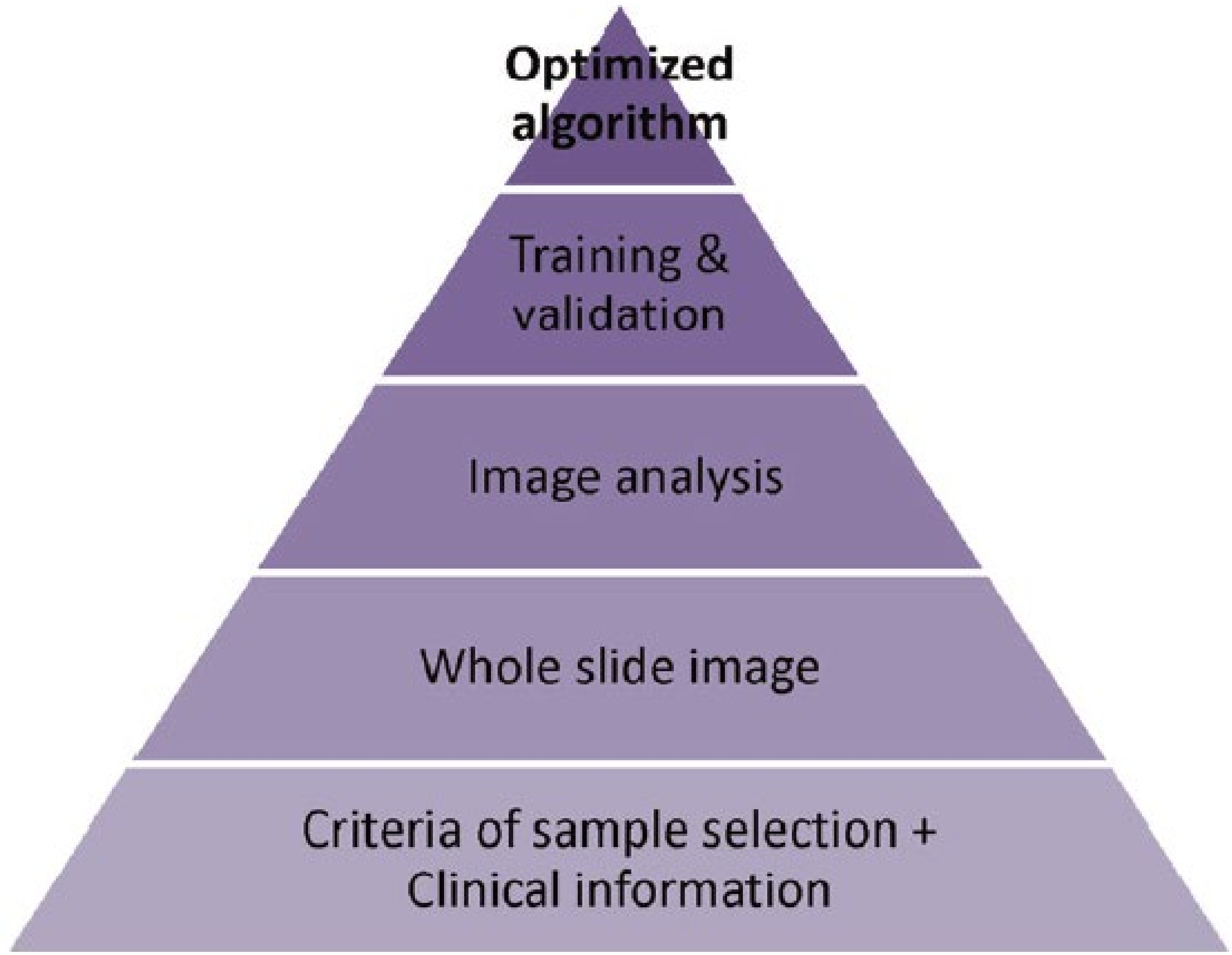
Laboratory Investigation (2021) 101:412–422

The principle of convolutional neural networks

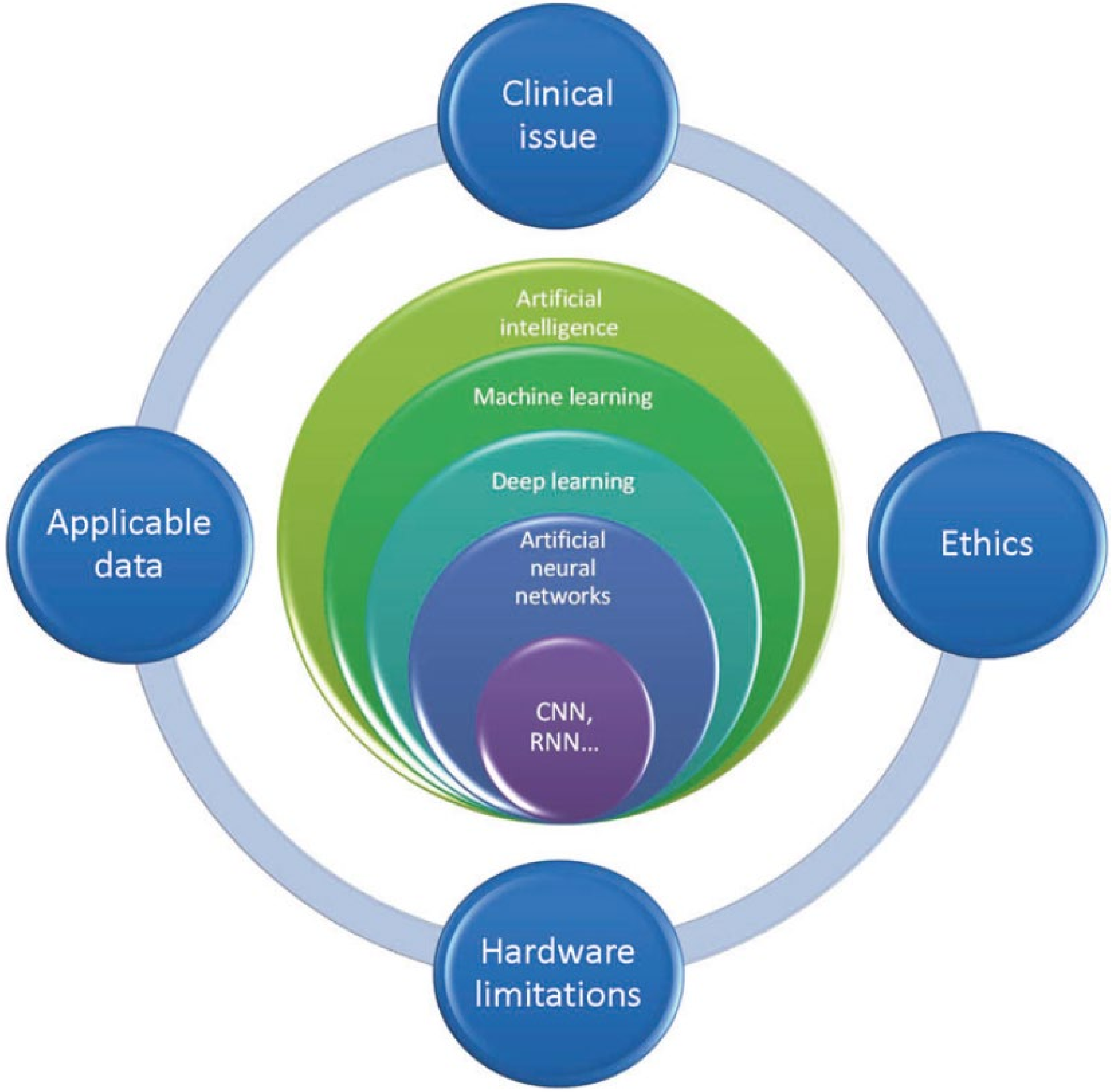


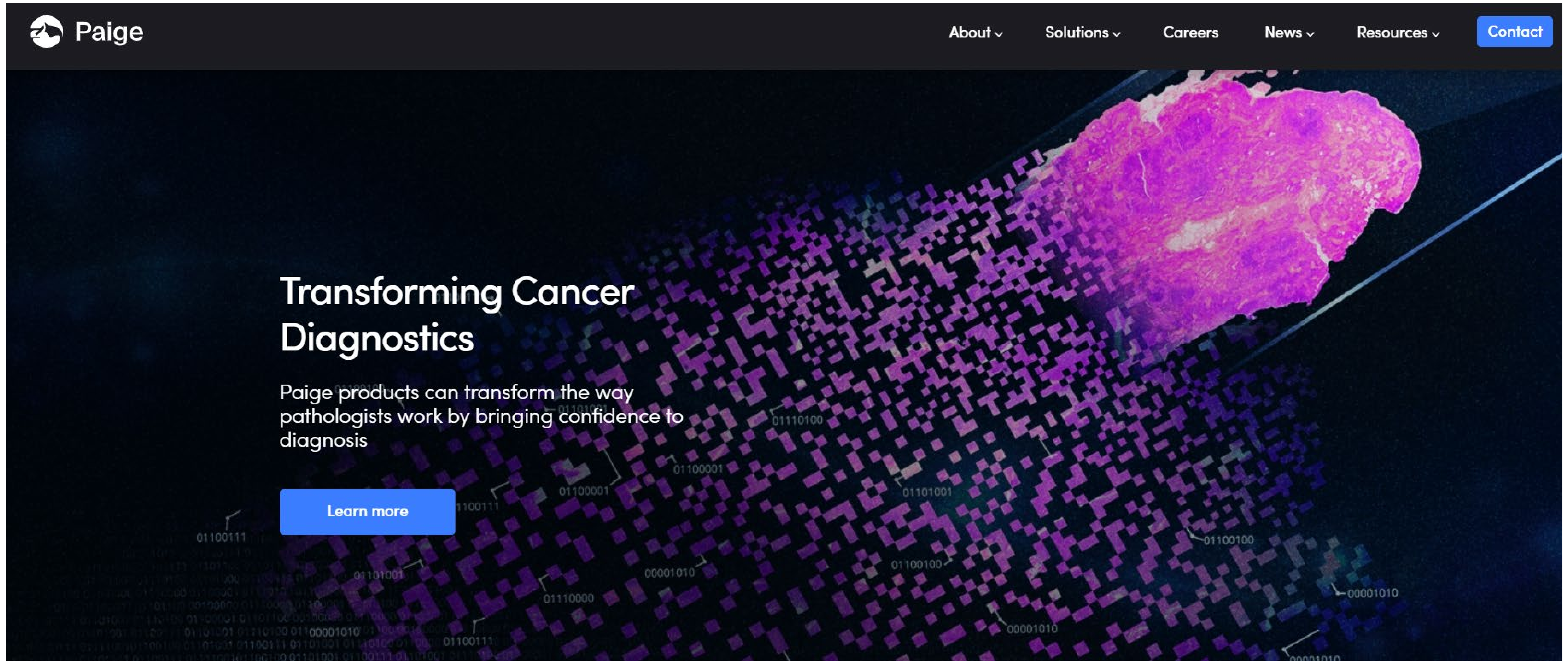
Laboratory Investigation (2021) 101:412–422

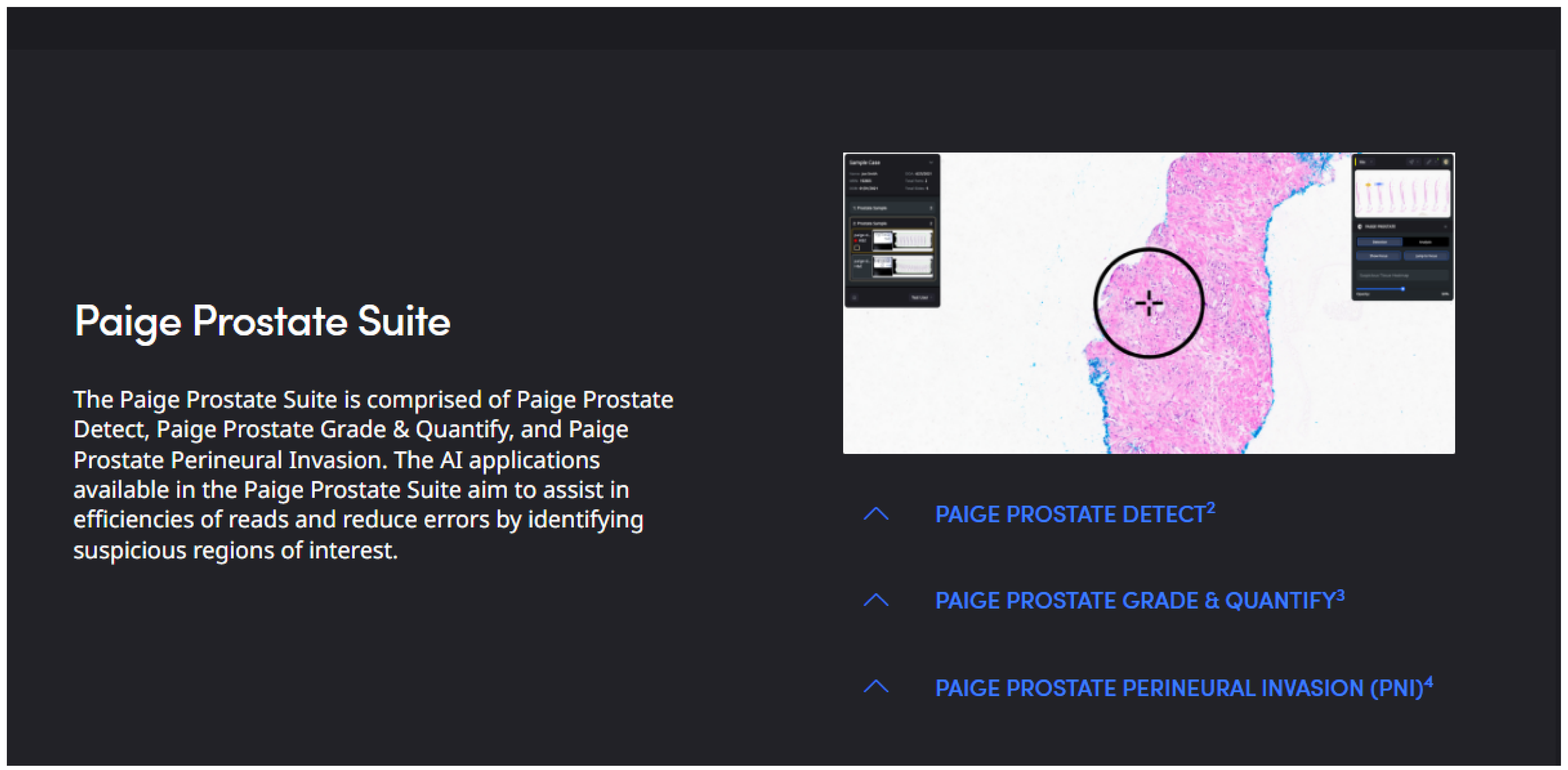
Flow chart of algorithmic training

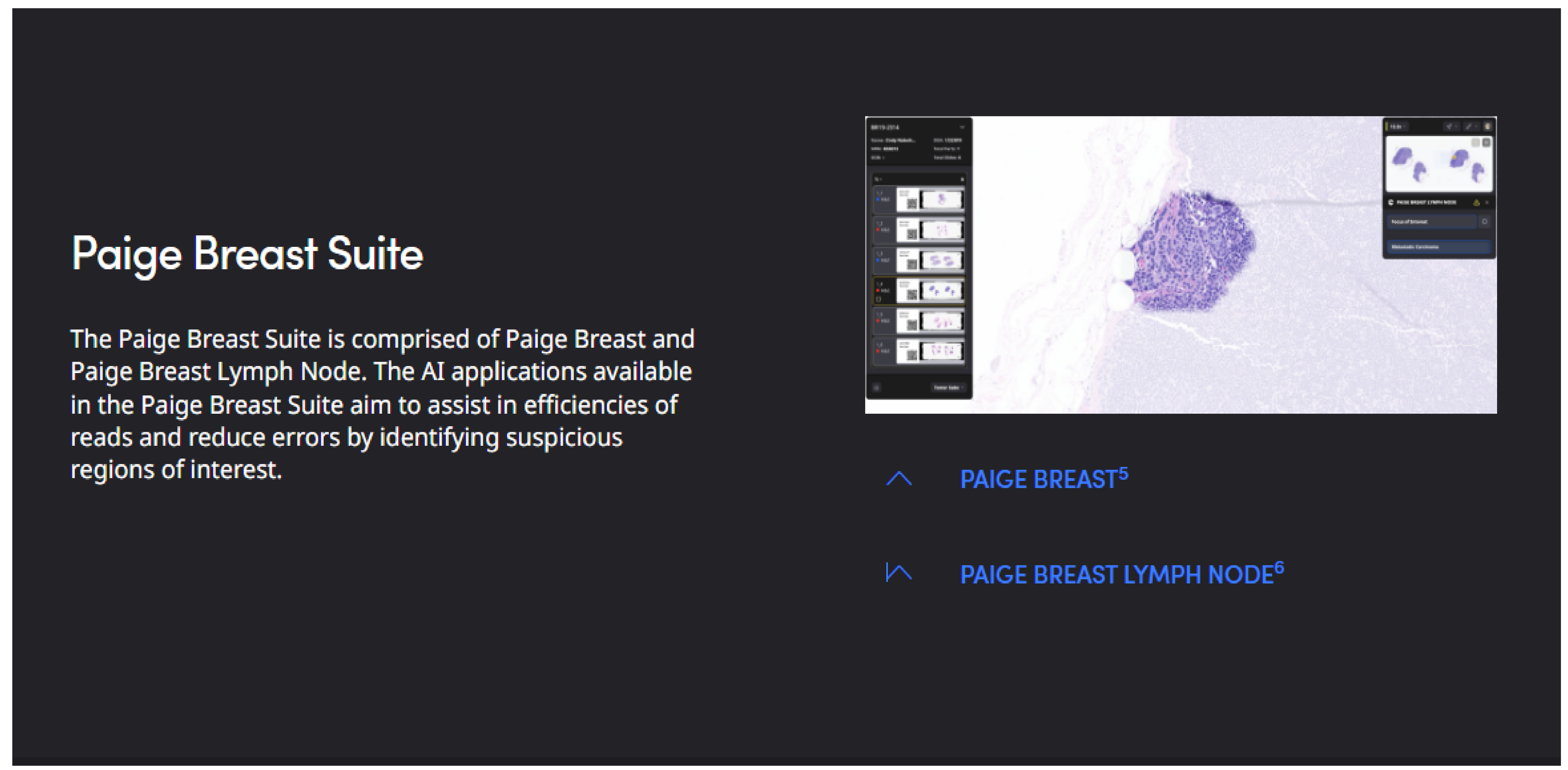


Laboratory Investigation (2021) 101:412–422


Clinical issue
Applicable data
Ethics
Hardware limitations
Artificial intelligence
Machine learning
Deep learning
Artificial neural networks
CNN, RNN...


Paige
About
Solutions
Careers
News
Resources
Contact
Transforming Cancer Diagnostics
Paige products can transform the way pathologists work by bringing confidence to diagnosis
Learn more


Paige Prostate Suite
The Paige Prostate Suite is comprised of Paige Prostate Detect, Paige Prostate Grade & Quantify, and Paige Prostate Perineural Invasion. The AI applications available in the Paige Prostate Suite aim to assist in efficiencies of reads and reduce errors by identifying suspicious regions of interest.
PAIGE PROSTATE DETECT[2]
PAIGE PROSTATE GRADE & QUANTIFY[3]
PAIGE PROSTATE PERINEURAL INVASION (PNI)[4]

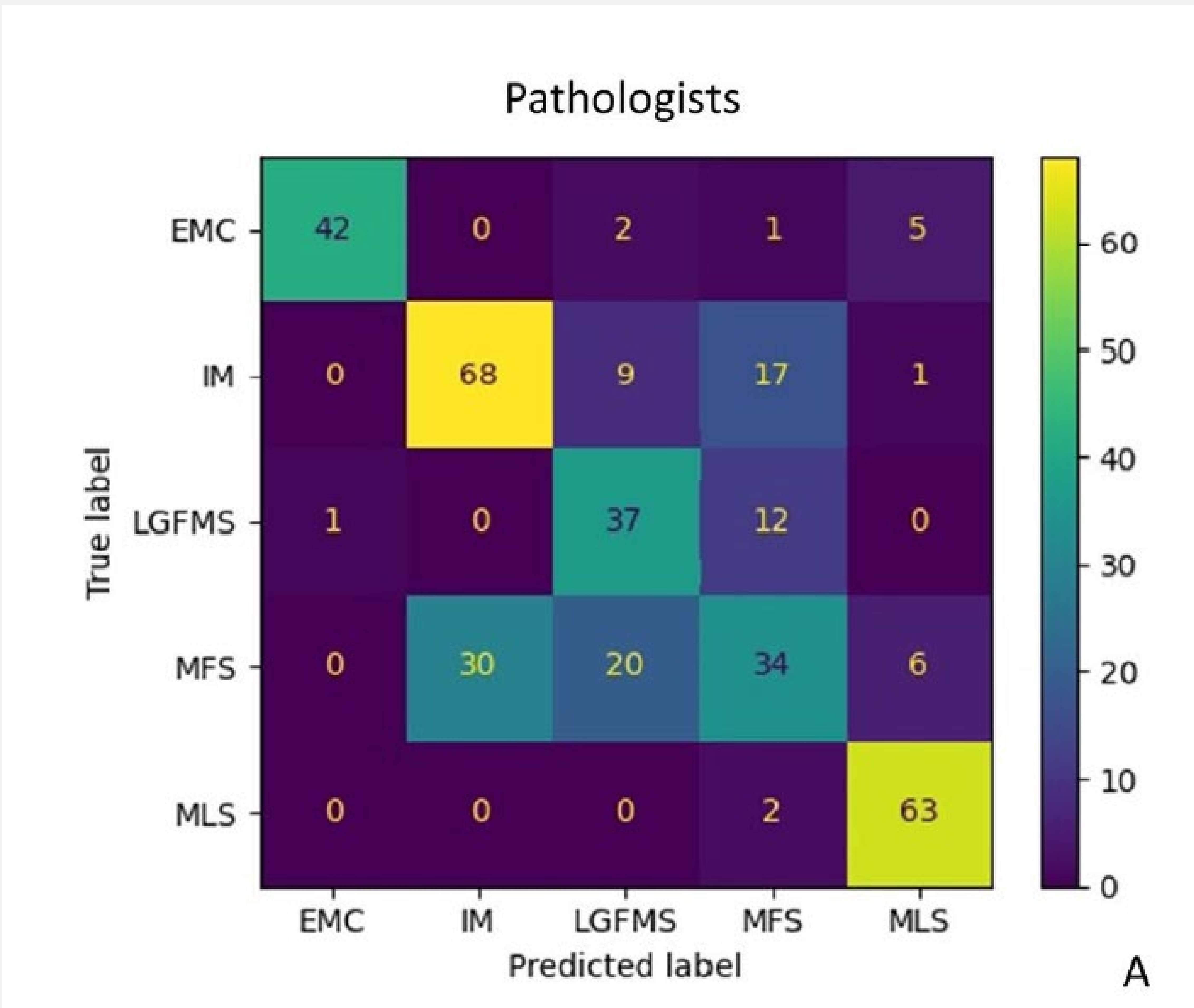

Paige Breast Suite
The Paige Breast Suite is comprised of Paige Breast and Paige Breast Lymph Node. The AI applications available in the Paige Breast Suite aim to assist in efficiencies of reads and reduce errors by identifying suspicious regions of interest.
PAIGE BREAST[5]
PAIGE BREAST LYMPH NODE[6]

Artificial intelligence significantly improves the diagnostic accuracy of deep myxoid soft tissue lesions in histology

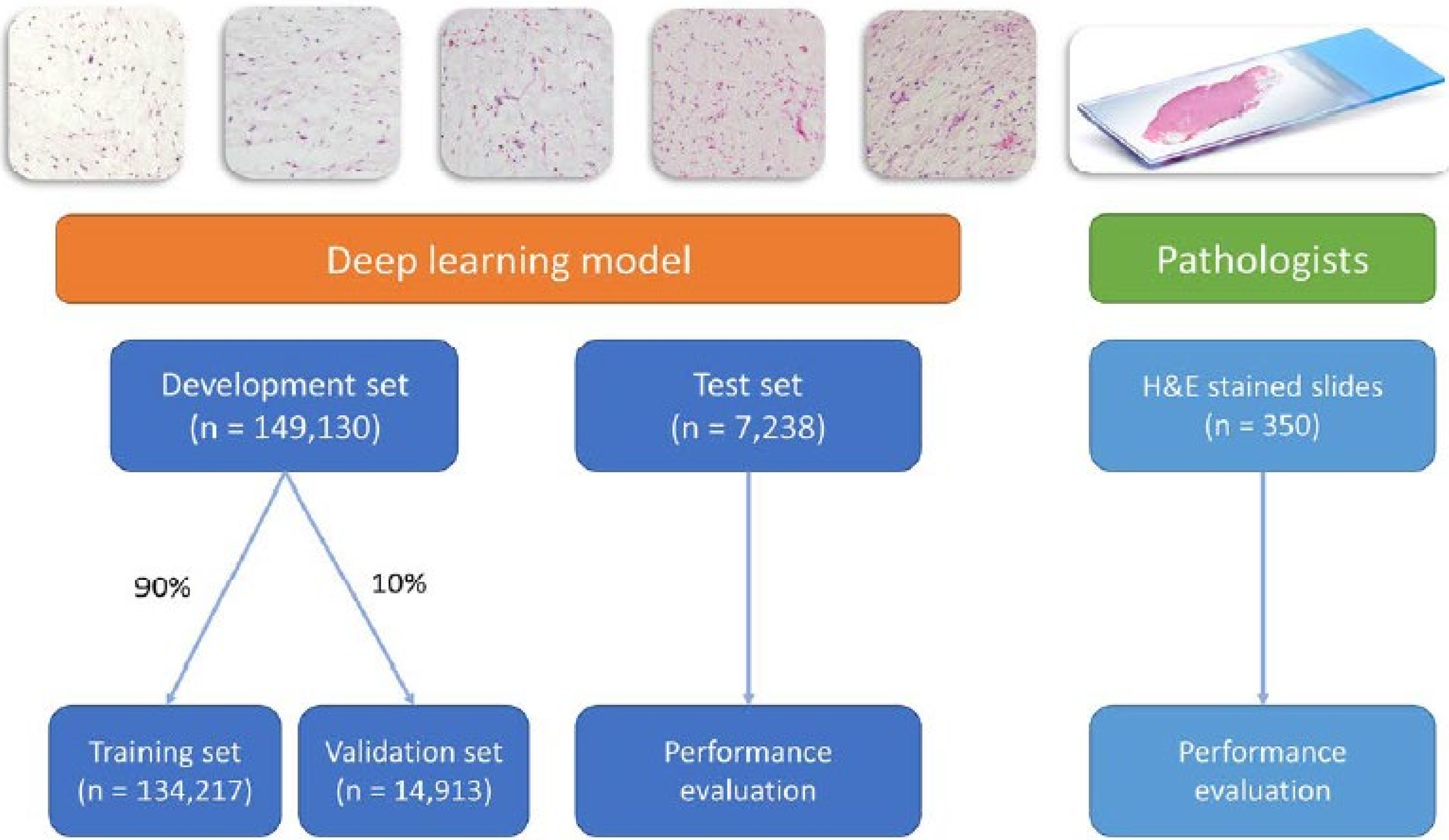
nature portfolio

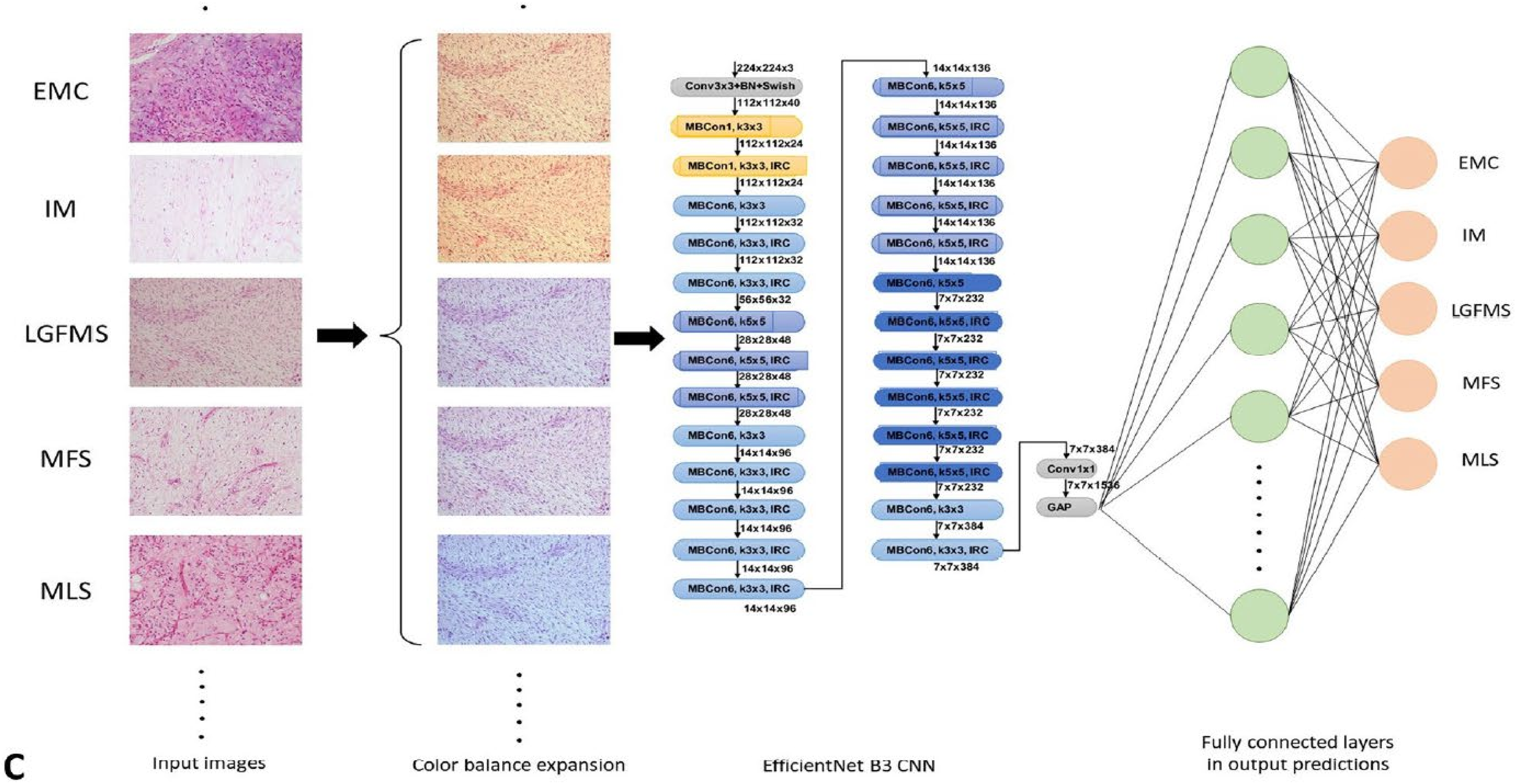
Maximus C. F. Yeung[1]✉ & Ivy S. Y. Cheng[2]

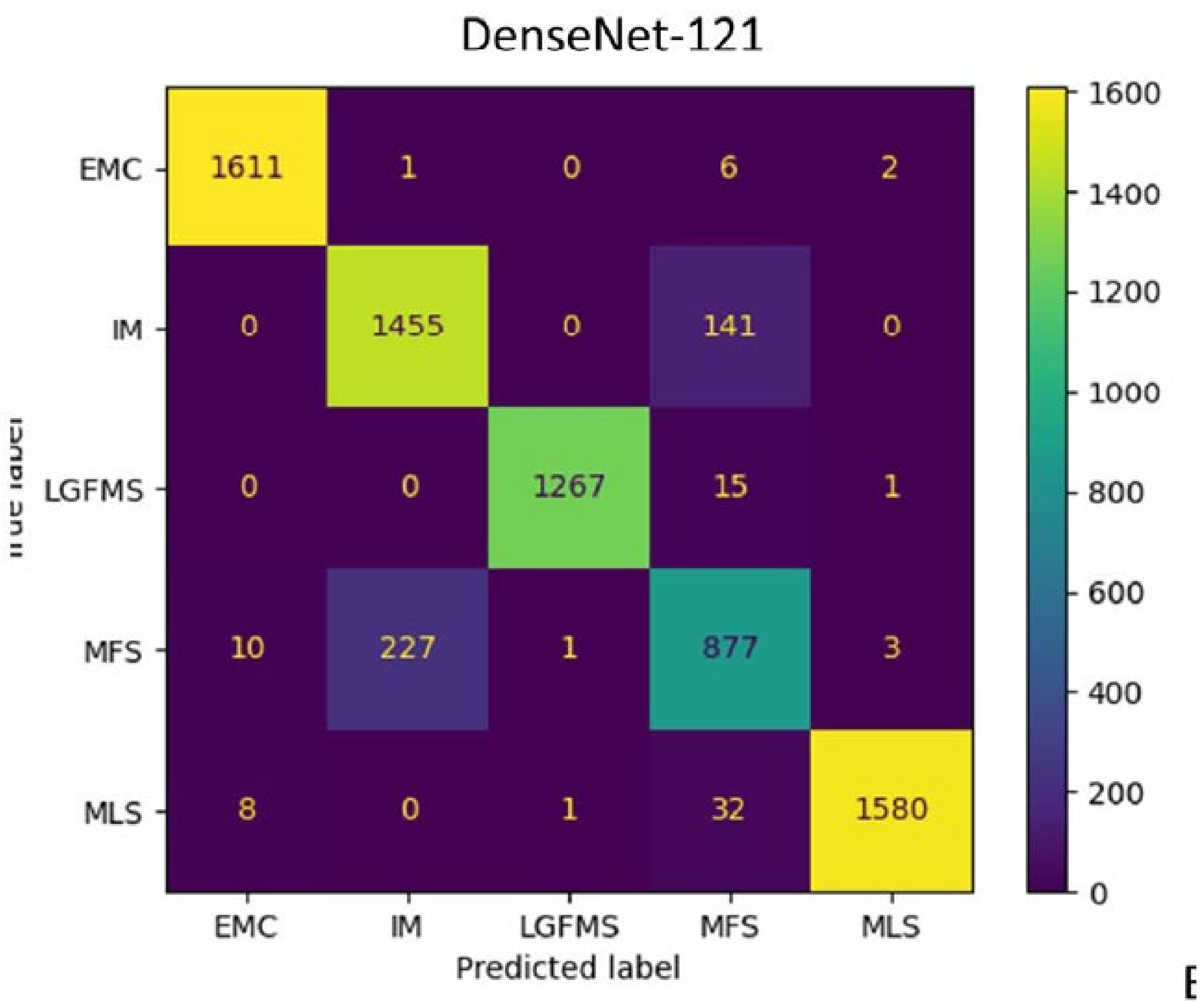
Scientific Reports | (2022) 12:6965

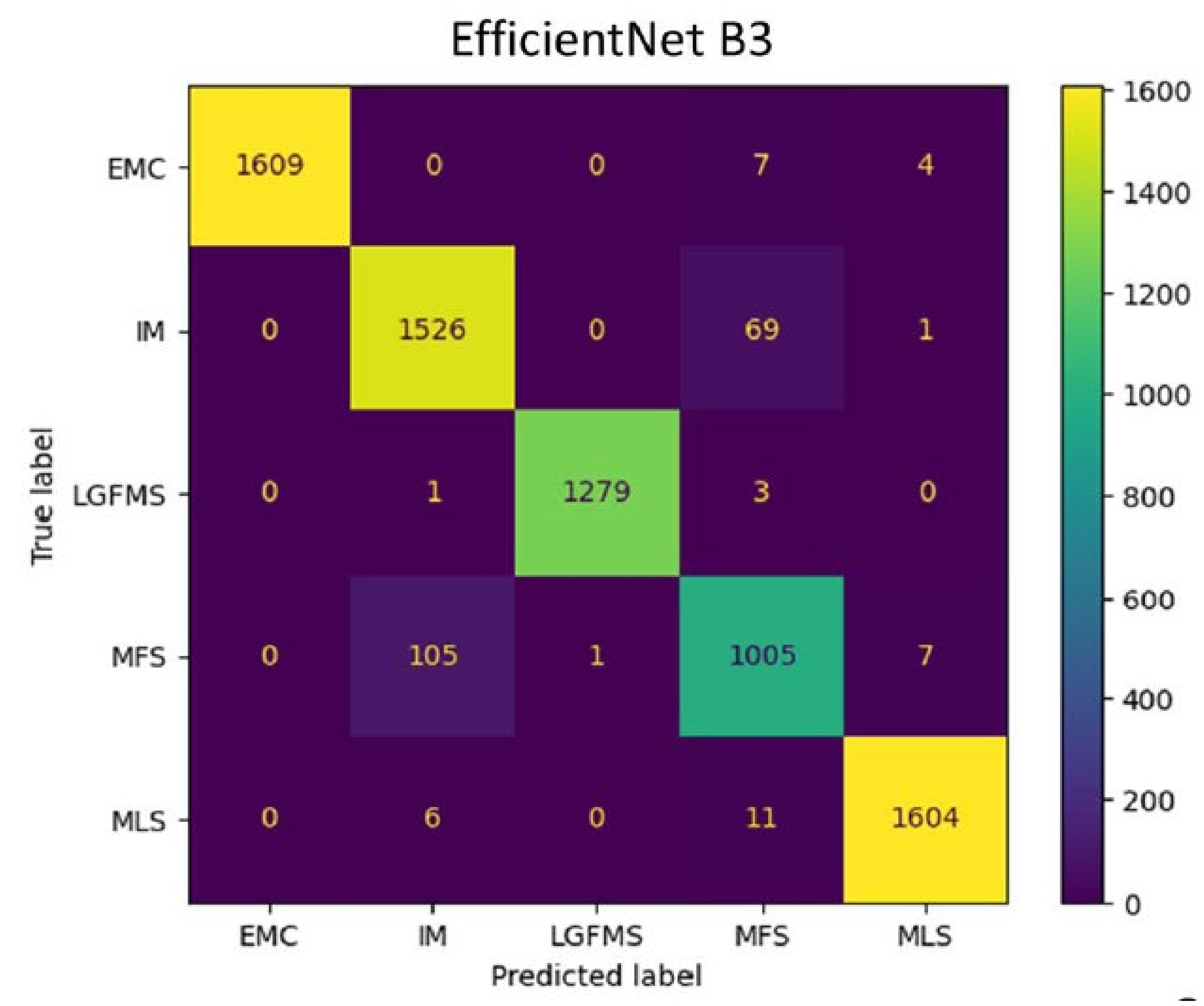


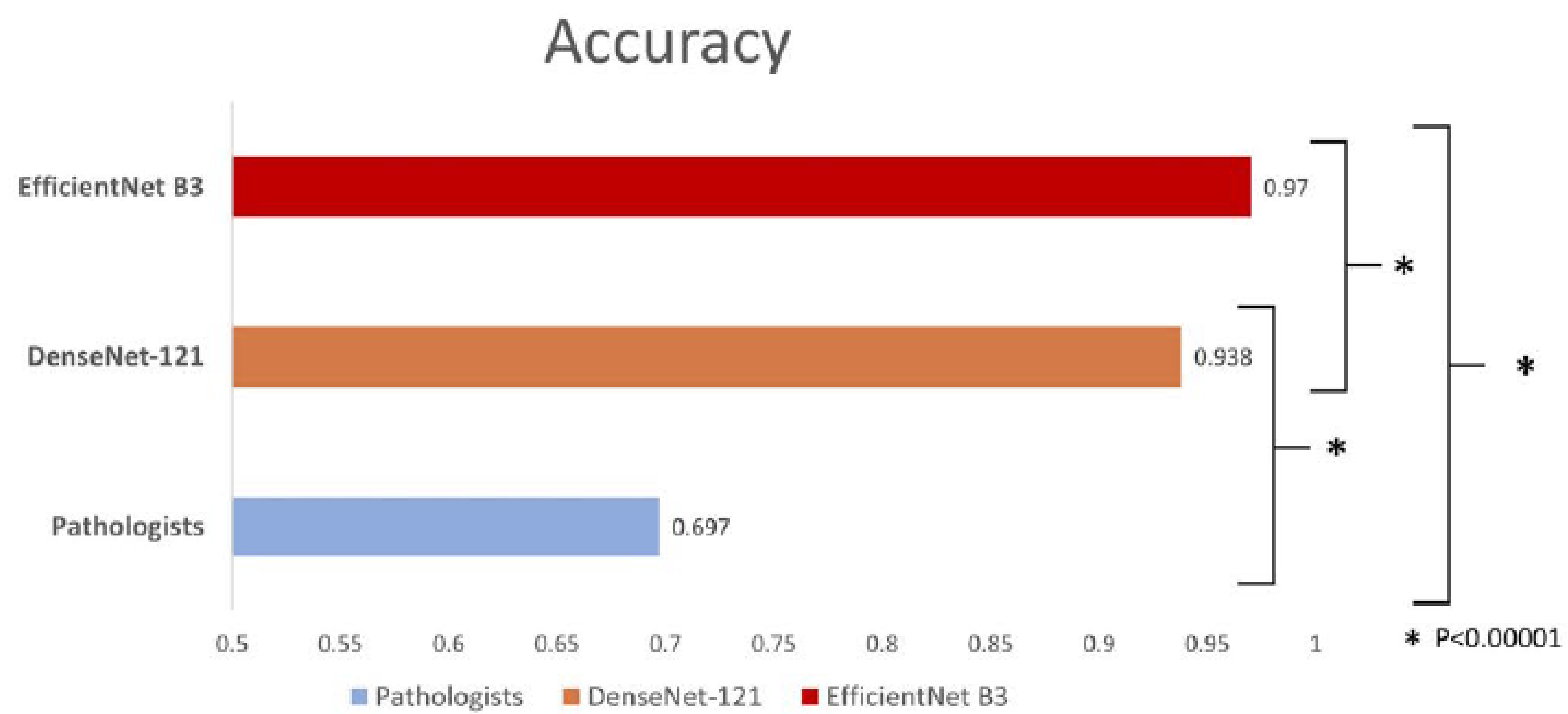
- 350 H&E slides from biopsies and excision specimens
  - Intramuscular myxoma
  - Myxofibrosarcoma
  - myxoid liposarcoma
  - low-grade fibromyxoid sarcoma
  - extraskeletal myxoid chondrosarcoma
- Read by 5 pathologists with different years of post-fellowship experience
- **Overall accuracy was only 69.7%**
- **Worse in biopsy samples 63.2%**

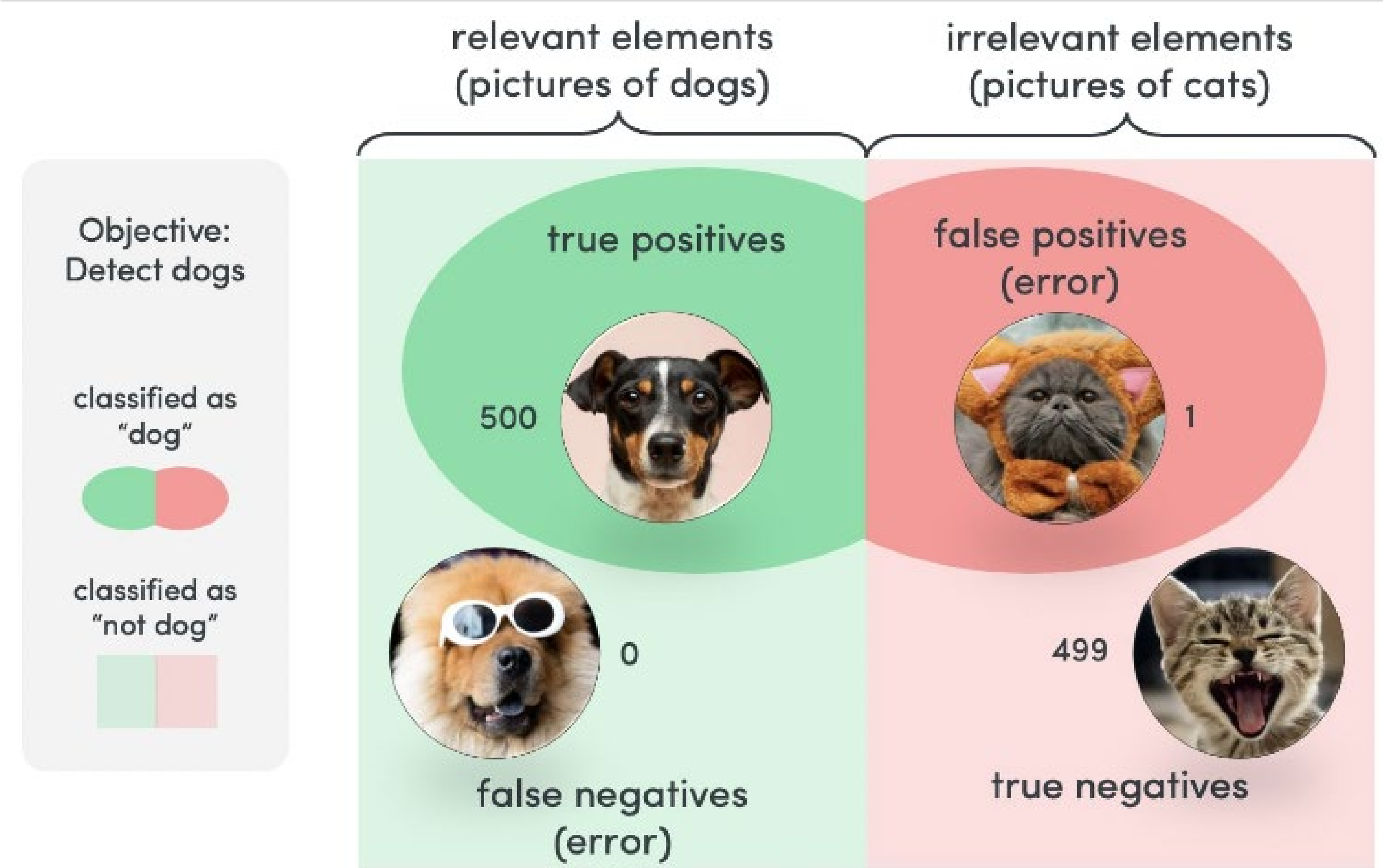

Deep learning model
Pathologists
Development set
(n = 149,130)
Test set
(n = 7,238)
H&E stained slides
(n = 350)
90%
10%
Training set
(n = 134,217)
Validation set
(n = 14,913)
Performance
evaluation
Performance
evaluation


EMC
IM
LGFMS
MFS
MLS
224x224x3
Conv3x3+BN+Swish
112x112x40
MBCon1, k3x3
112x112x24
MBCon1, k3x3, IRC
MBCon6, k3x3
112x112x32
MBCon6, k3x3, IRC
56x56x32
MBCon6, k5x5
28x28x48
MBCon6, k5x5, IRC
MBCon6, k3x3
14x14x96
MBCon6, k3x3, IRC
14x14x136
MBCon6, k5x5
MBCon6, k5x5, IRC
7x7x232
MBCon6, k3x3
7x7x384
MBCon6, k3x3, IRC
Conv1x1
7x7x1536
GAP
EMC
IM
LGFMS
MFS
MLS
C
Input images
Color balance expansion
EfficientNet B3 CNN
Fully connected layers
in output predictions

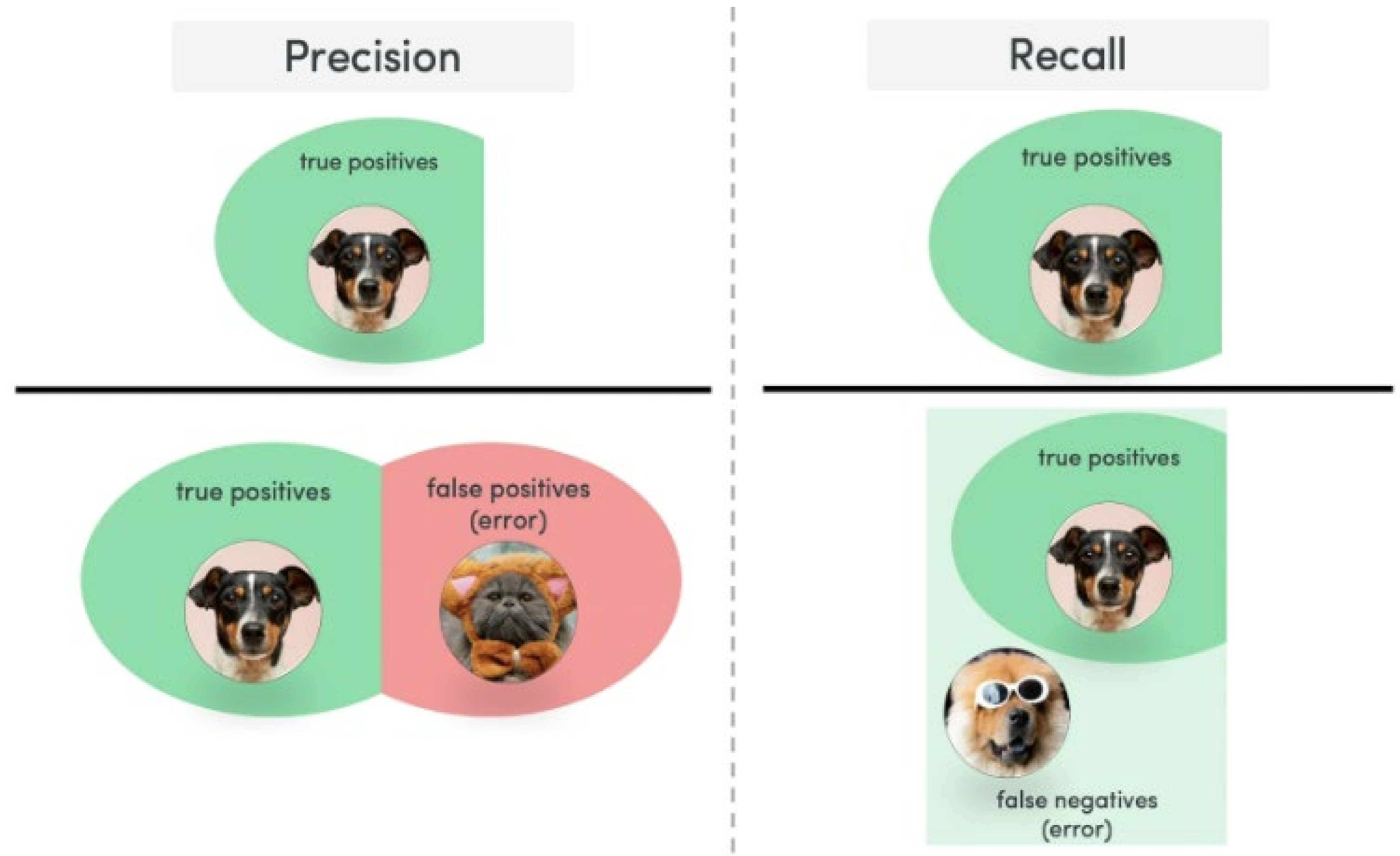

DenseNet-121
EMC
IM
LGFMS
MFS
MLS
1611 1 0 6 2
0 1455 0 141 0
0 0 1267 15 1
10 227 1 877 3
8 0 1 32 1580
Predicted label
0
200
400
600
800
1000
1200
1400
1600


EfficientNet B3
True label
EMC
IM
LGFMS
MFS
MLS
1609 0 0 7 4
0 1526 0 69 1
0 1 1279 3 0
0 105 1 1005 7
0 6 0 11 1604
Predicted label
0
200
400
600
800
1000
1200
1400
1600


Accuracy
EfficientNet B3
DenseNet-121
Pathologists
0.97
0.938
0.697
*
*
*
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1
* P<0.00001
Pathologists
DenseNet-121
EfficientNet B3



Matrix of the choices the dog/cat image classifier made



How precision and recall are calculated

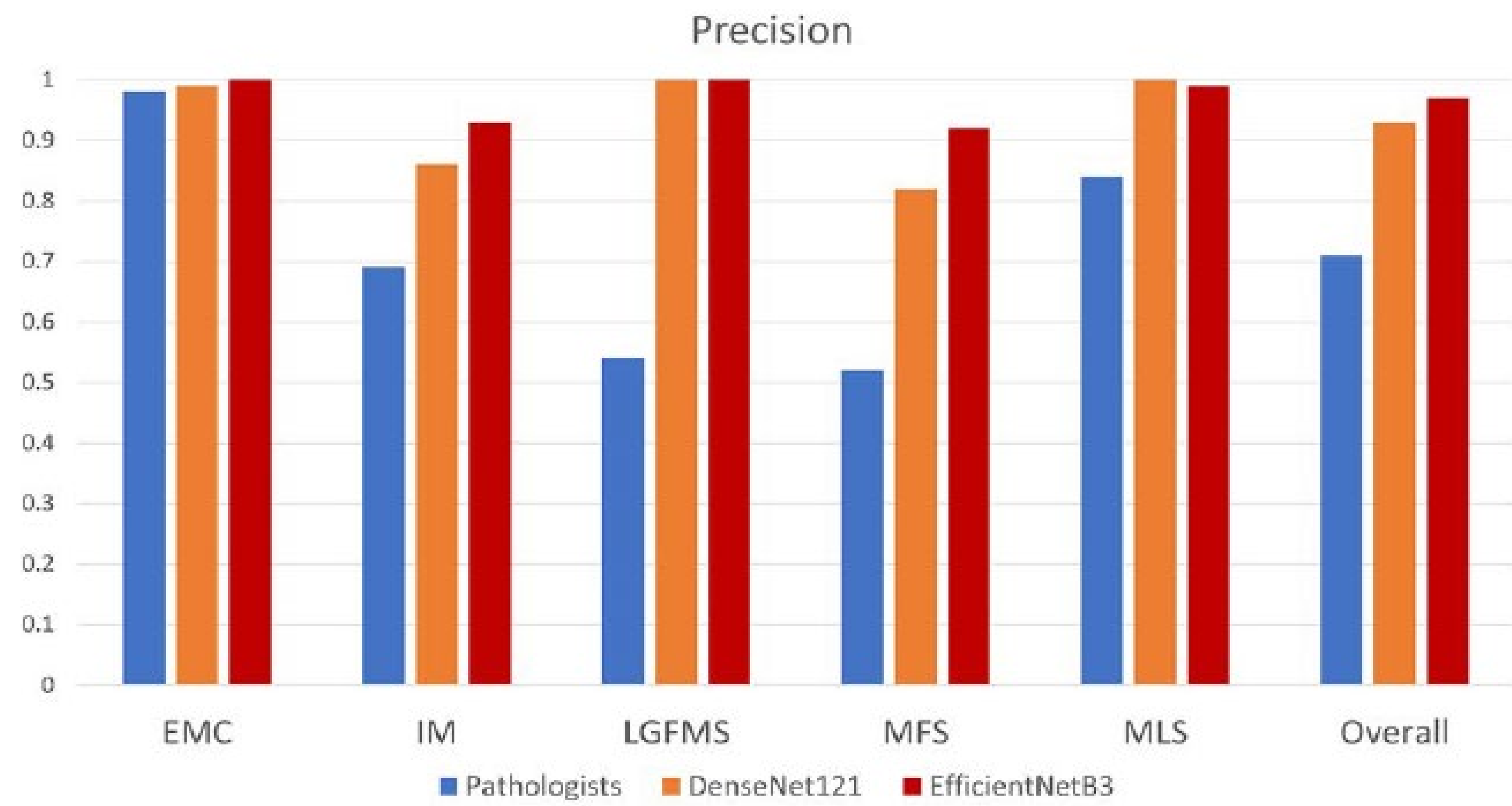
Precision

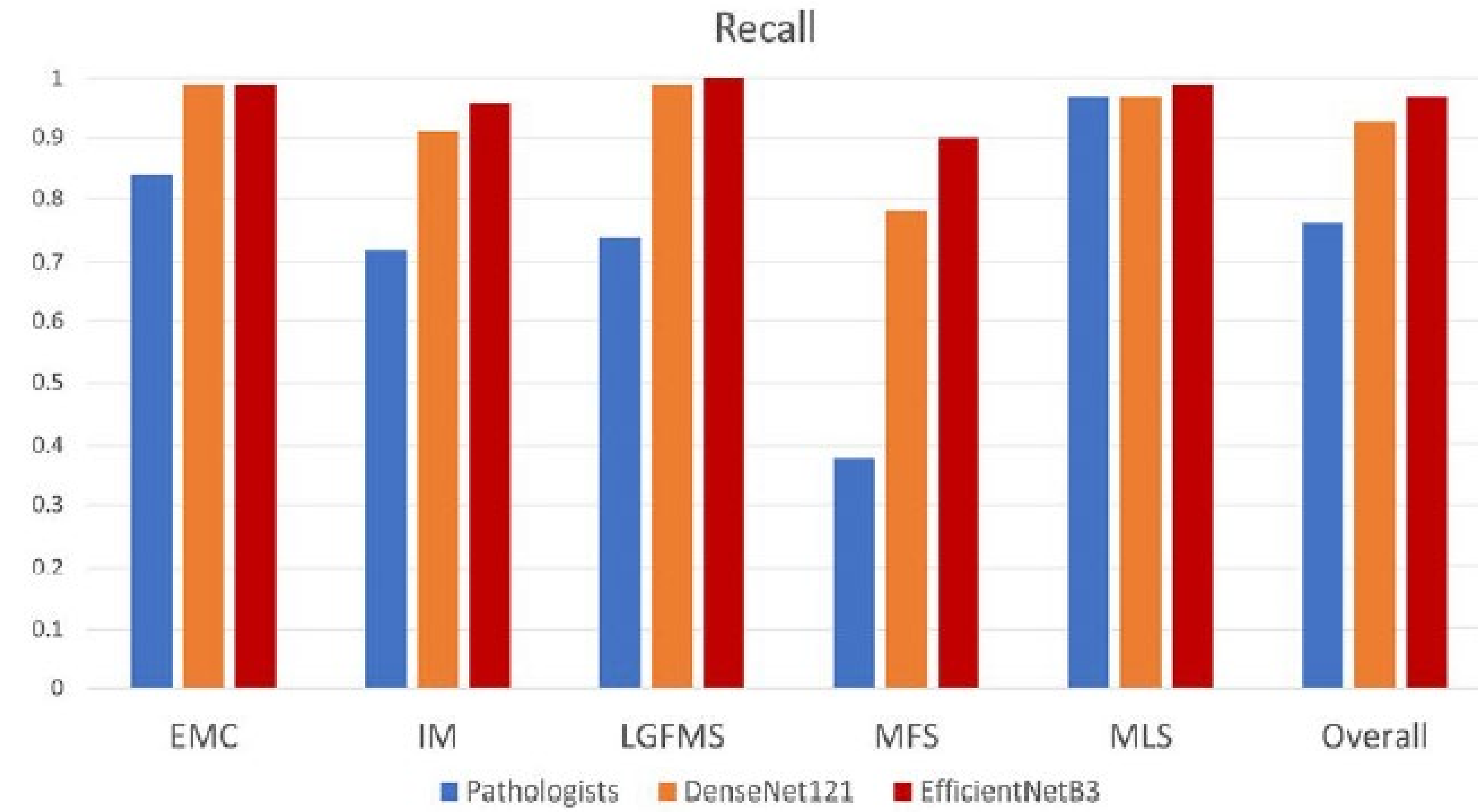
The precision of a model describes how many detected items are truly relevant. It is calculated by dividing the true positives by overall positives.

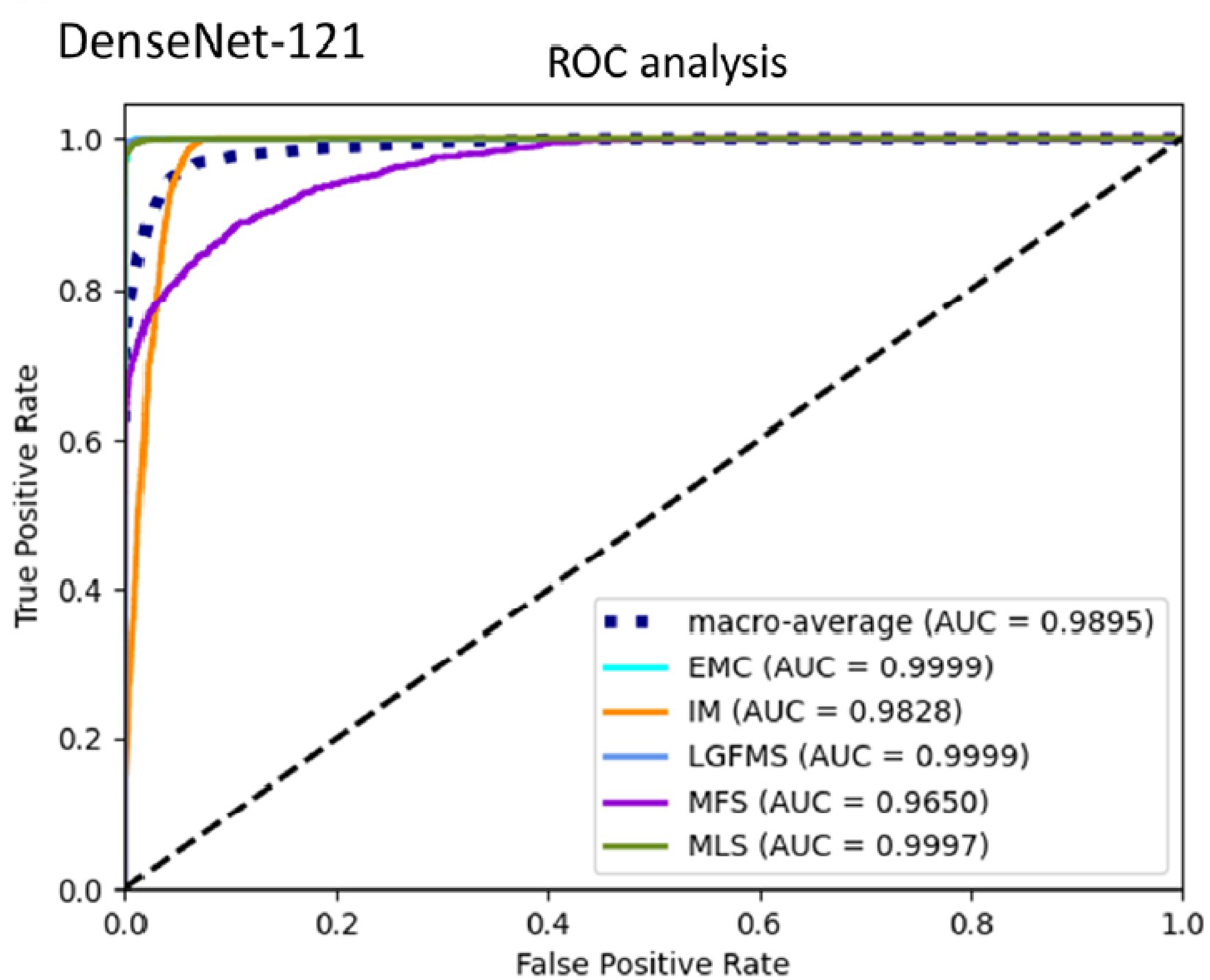
Recall

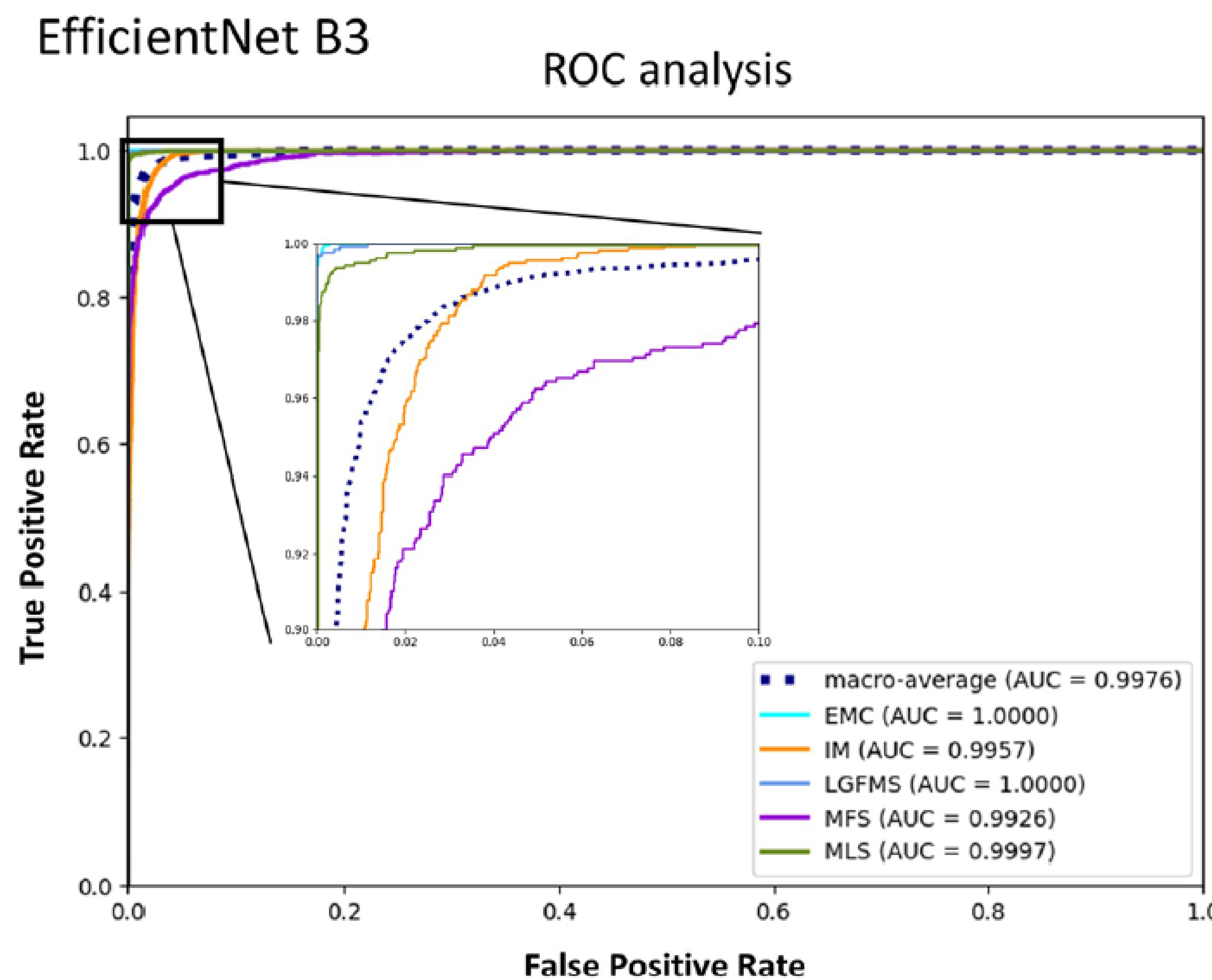
Recall is a measure of how many relevant elements were detected. Therefore it divides true positives by the number of relevant elements.

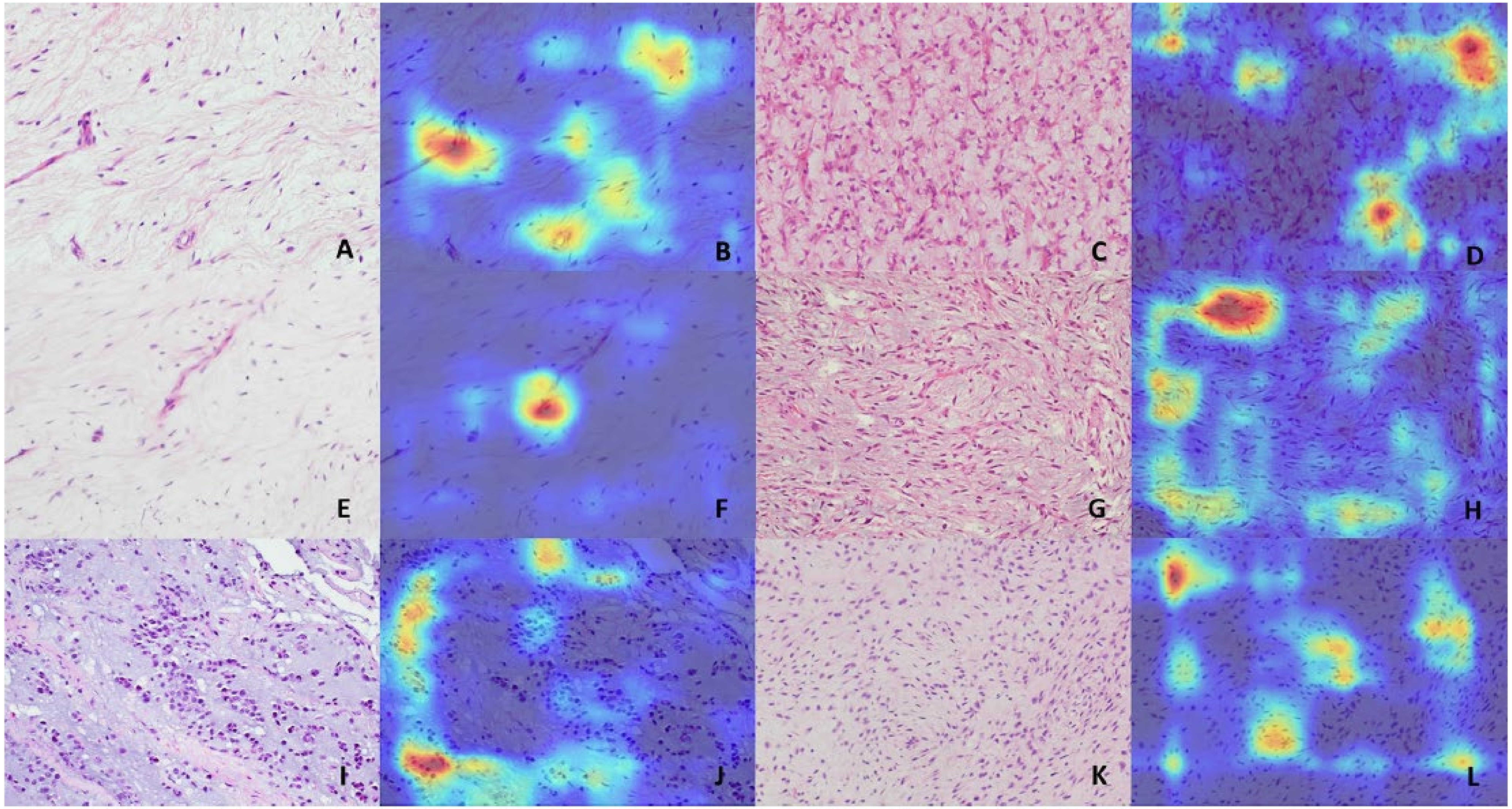
Combining precision and recall: The F-measure


Precision
EMC
IM
LGFMS
MFS
MLS
Overall
Pathologists
DenseNet121
EfficientNetB3


Recall
EMC
IM
LGFMS
MFS
MLS
Overall
Pathologists
DenseNet121
EfficientNetB3


DenseNet-121
ROC analysis
True Positive Rate
False Positive Rate
macro-average (AUC = 0.9895)
EMC (AUC = 0.9999)
IM (AUC = 0.9828)
LGFMS (AUC = 0.9999)
MFS (AUC = 0.9650)
MLS (AUC = 0.9997)


EfficientNet B3
ROC analysis
True Positive Rate
False Positive Rate
macro-average (AUC = 0.9976)
EMC (AUC = 1.0000)
IM (AUC = 0.9957)
LGFMS (AUC = 1.0000)
MFS (AUC = 0.9926)
MLS (AUC = 0.9997)


A
B
C
D
E
F
G
H
I
J
K
L

- Artificial intelligence (AI) using deep learning model with convolutional neural network **outperformed pathologist in diagnosis deep myxoid lesions**

- Potential for AI to augment pathologist’s eye with information that cannot be done by human examination

- There remains a small number of cases, especially between **IM and low-grade MFS**, where these two entities are truly undifferentiable even with AI

The American Journal of Pathology, Vol. 192, No. 6, June 2022

ELSEVIER

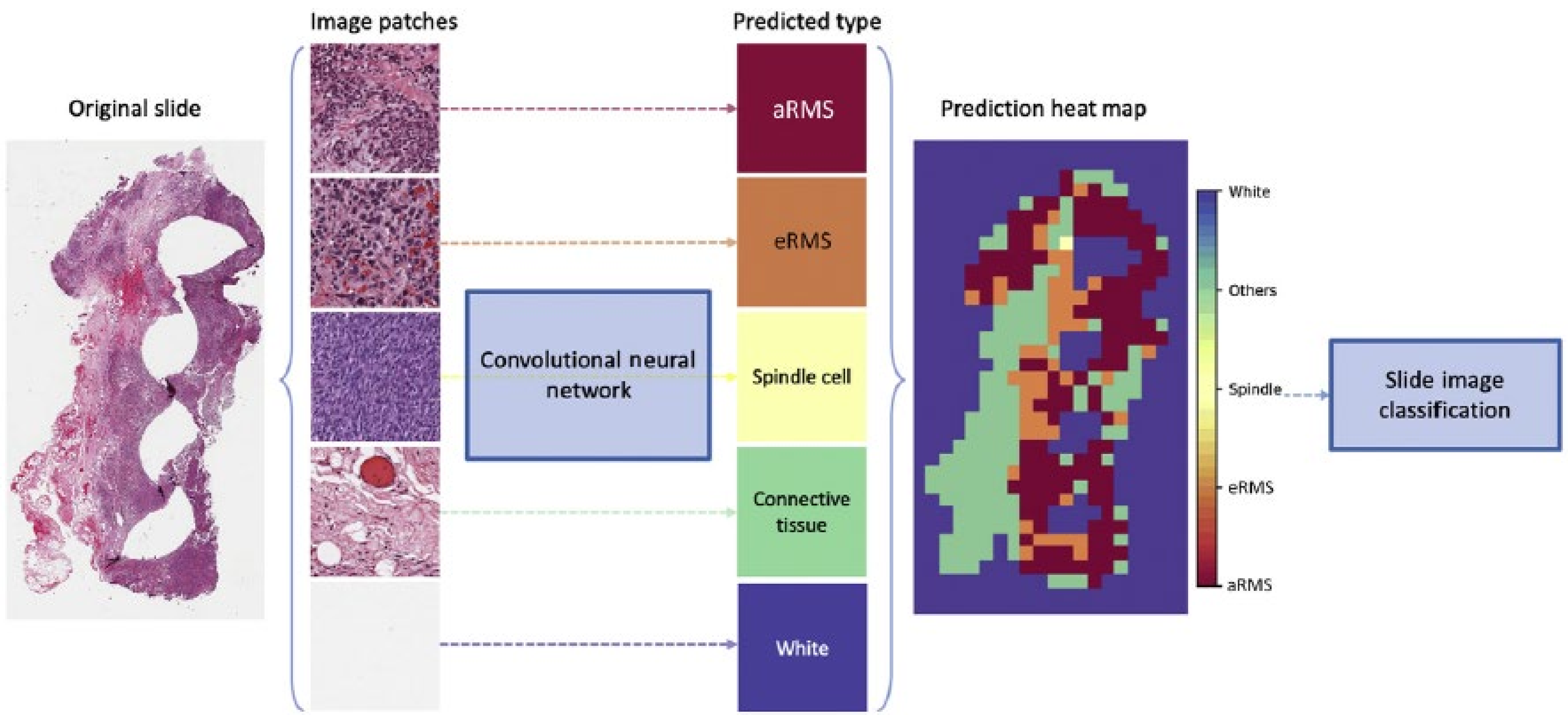
The American Journal of
PATHOLOGY

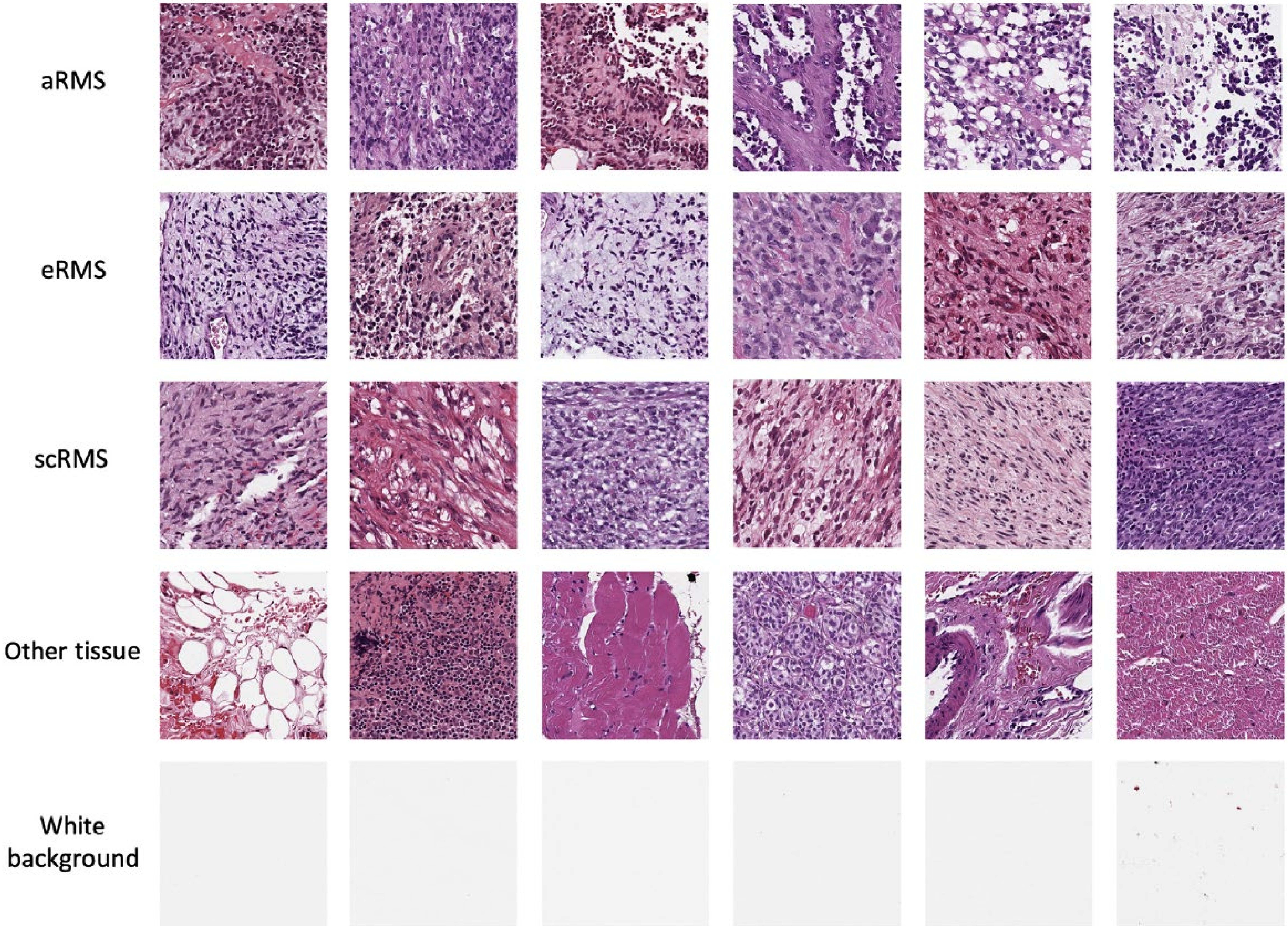
ajp.amjpathol.org

MACHINE LEARNING, COMPUTATIONAL PATHOLOGY, AND BIOPHYSICAL IMAGING

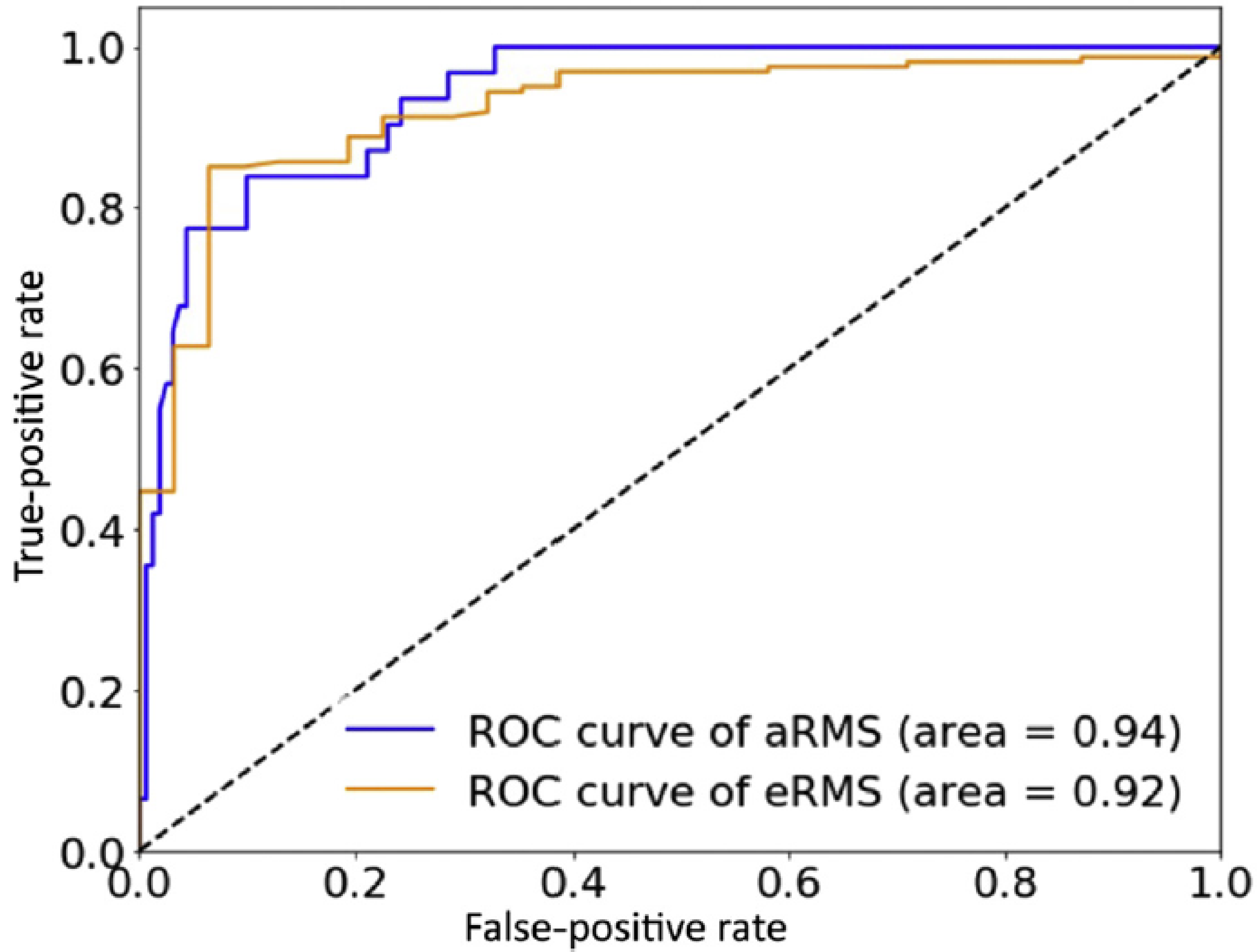
Deep Learning of Rhabdomyosarcoma Pathology Images for Classification and Survival Outcome Prediction

Xinyi Zhang,[*] Shidan Wang,[*] Erin R. Rudzinski,[†] Saloni Agarwal,[‡] Ruichen Rong,[*] Donald A. Barkauskas,[§] Ovidiu Daescu,[‡] Lauren Furman Cline,[¶] Rajkumar Venkatramani,[||**] Yang Xie,[*††] Guanghua Xiao,[*††] and Patrick Leavey[¶]




aRMS
eRMS
scRMS
Other tissue
White background


345 RMS slides
272 Slides with pathology annotation
80 Slides used to develop patch-level RMS classification model
192 Slides reserved to test the RMS classification model on slide level
n of Slides | aRMS | eRMS | scRMS
Training set | 25 | 25 | 6
Validation set | 5 | 5 | 2
Test set | 5 | 5 | 2
Total | 35 | 35 | 10
Extract image patches
n of patches | aRMS | eRMS | scRMS | Other tissue | White space
Training set | 3107 | 3793 | 1246 | 1174 | 1338
Validation set | 876 | 570 | 200 | 291 | 432
Test | 513 | 368 | 203 | 289 | 301
Train
Test
Test data set
Patch-level RMS classification model
Prediction results
Subtype probability heat map
(Examples are shown in Supplemental Figure S2)
Summarize
Whole slide level subtype classification
Figure 3. Example image patches
Supplemental Figure S1. Patch-level classification model confusion matrix
Figure 4. ROC curve slide-level subtype classification
Take the major subtype
Transfer learning
196 eRMS slides
136 Slides to test the prognosis model
60 Slides used to develop a patch-level eRMS prognostic model
Test data set
Patch-level eRMS prognosis model
n of Slides | Low risk | High risk
Training set | 24 | 24
Validation set | 6 | 6
Total | 30 | 30
Extract image patches
Prediction results
eRMS prognosis heat map
Summarize
Train
n of patches | Low risk | High risk
Training set | 1337 | 1146
Validation set | 387 | 204
Total | 1724 | 1350
Patient-level prognosis prediction
Figure 5. KM curve for patient prognosis



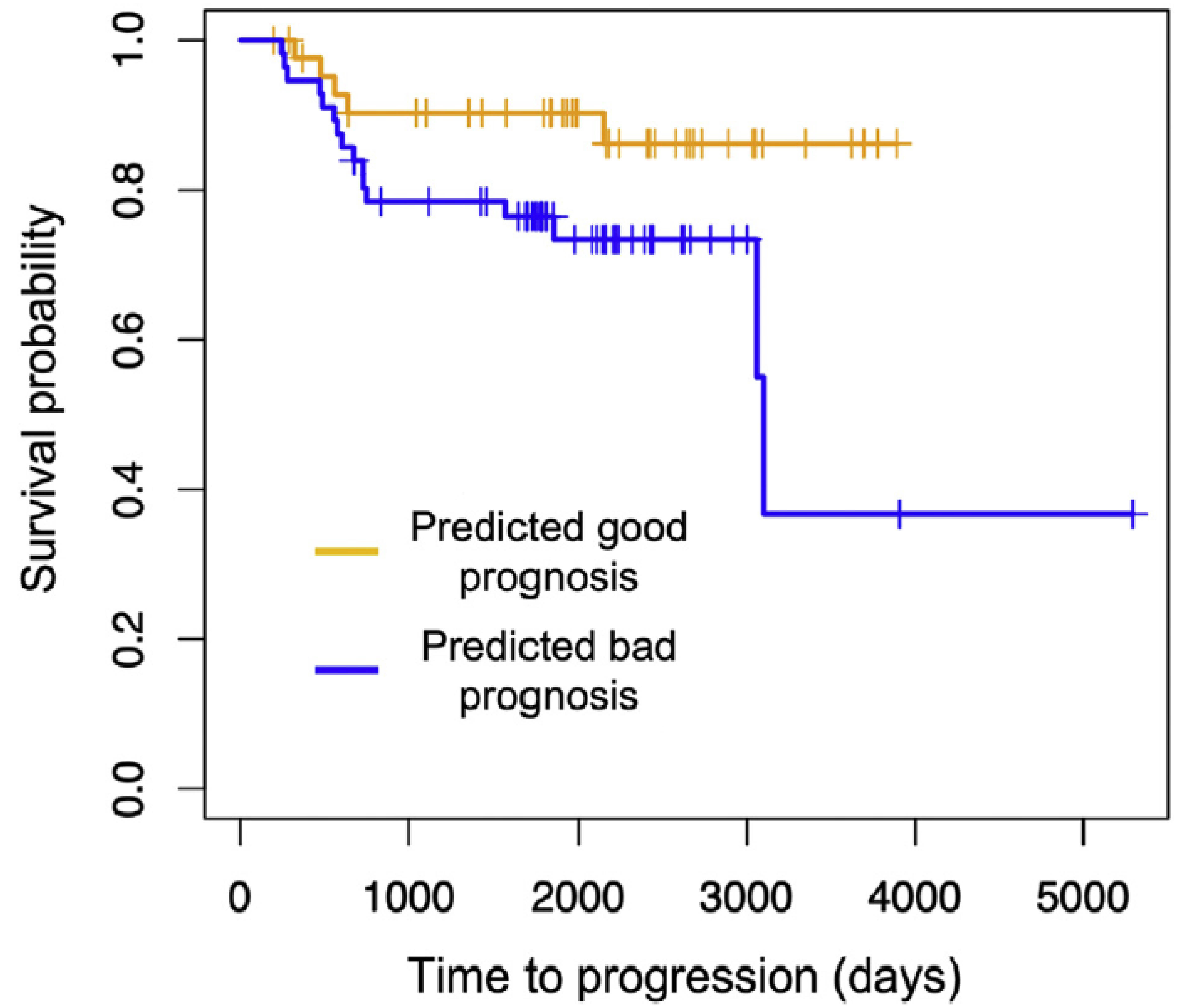


Table 2 Multivariate Analysis to Adjust Image-Based Prognostic Model with Clinical Variables Using Cox Proportional Hazard Model

Clinical variables	HR	Lower 95% CI	Upper 95% CI	*P* value
Image-based prognostic model	4.64	1.05	20.57	0.04
Age	1.00	1.00	1.00	0.78
Sex (male)	1.04	0.37	2.93	0.94

USCAP

www.nature.com/modpathol

ARTICLE

Check for updates

Machine learning for rhabdomyosarcoma histopathology

Arthur O. Frankel[1,22], Melvin Lathara[2,22], Celine Y. Shaw[1], Owen Wogmon[1], Jacob M. Jackson[1], Mattie M. Clark[1], Navah Eshraghi[1], Stephanie E. Keenen[1], Andrew D. Woods[1], Reshma Purohit[1], Yukitomo Ishi[3], Nirupama Moran[4], Mariko Eguchi[5], Farhat Ul Ain Ahmed[6], Sara Khan[7], Maria Ioannou[8], Konstantinos Perivoliotis[9], Pin Li[10], Huixia Zhou[10], Ahmad Alkhaledi[11], Elizabeth J. Davis[12], Danielle Galipeau[13], R. L. Randall[14], Agnieszka Wozniak[15], Patrick Schoffski[15], Che-Jui Lee[15], Paul H. Huang[16], Robin L. Jones[16], Brian P. Rubin[17], Morgan Darrow[18], Ganapati Srinivasa[2], Erin R. Rudzinski[19], Sonja Chen[20,21]✉, Noah E. Berlow[1,23]✉ and Charles Keller[1,23]✉

Modern Pathology (2022) 35:1193–1203

Predicting Molecular Subtype and Survival of Rhabdomyosarcoma Patients using Deep Learning of H&E Images: A Report from the Children's Oncology Group

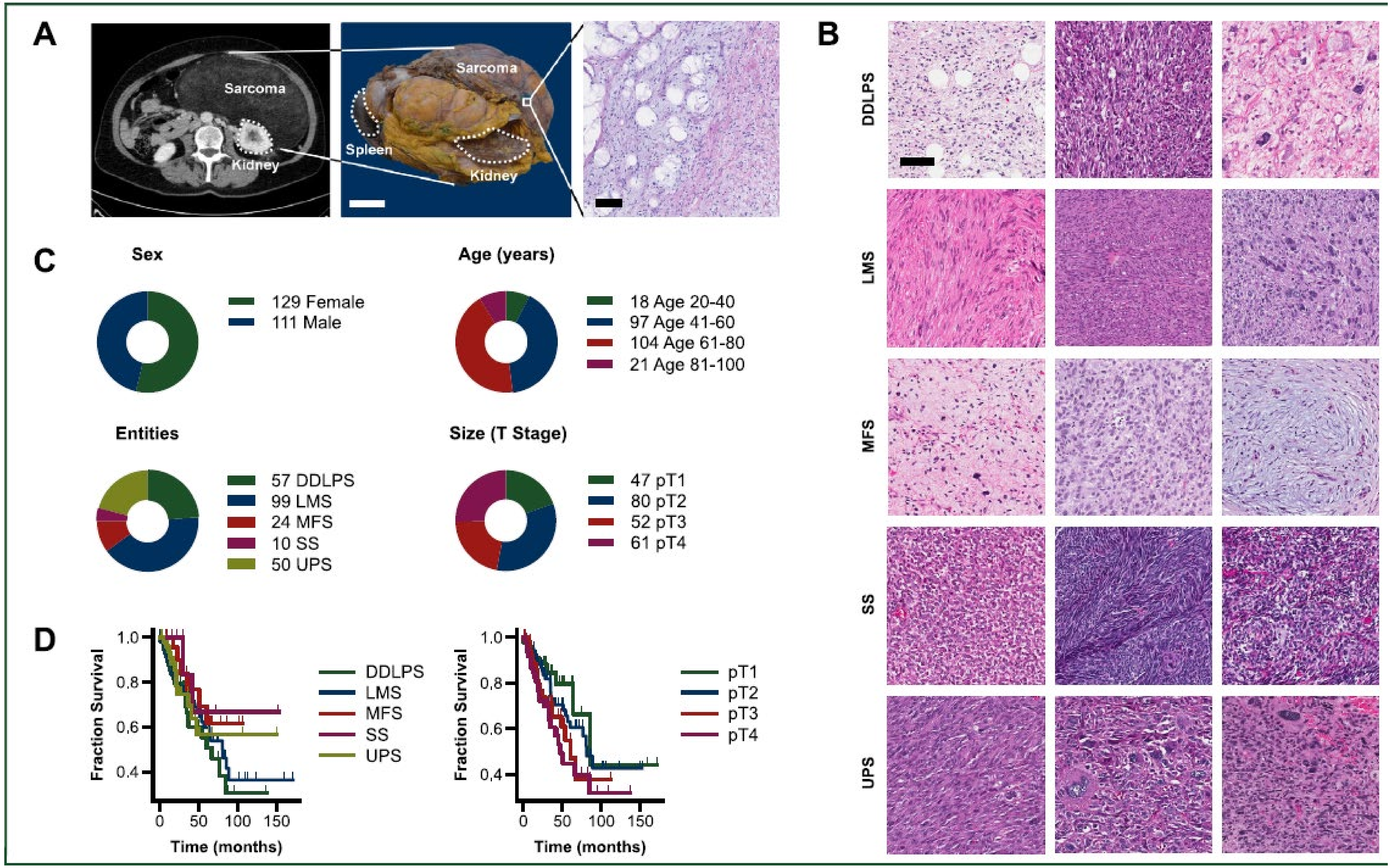
David Milewski[1#], Hyun Jung[2#], G. Thomas Brown[2,3#], Yanling Liu[2], Ben Somerville[1], Curtis Lisle[2,4], Marc Ladanyi[5], Erin R. Rudzinski[6], Hyoyoung Choo-Wosoba[7], Donald A. Barkauskas[8,9], Tammy Lo[9], David Hall[9], Corinne M. Linardic[10], Jun S. Wei[1], Hsien-Chao Chou[1], Stephen X. Skapek[11], Rajkumar Venkatramani[12], Peter K. Bode[13], Seth M. Steinberg[7], George Zaki[14], Igor B. Kuznetsov[15], Douglas S. Hawkins[16], Jack F. Shern[17], Jack Collins[2], and Javed Khan[1]*.

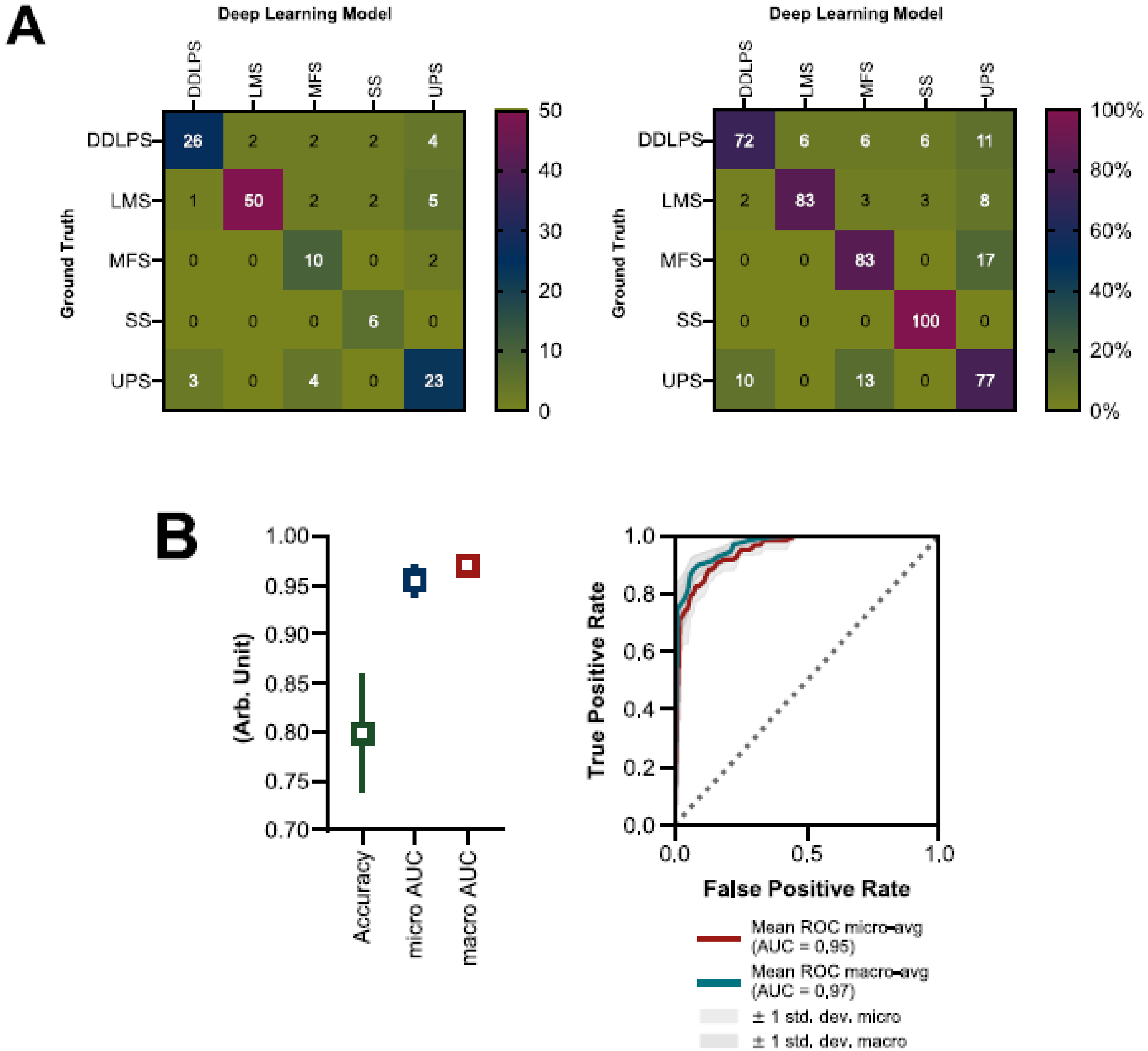
Deep learning for diagnosis and survival prediction in soft tissue sarcoma

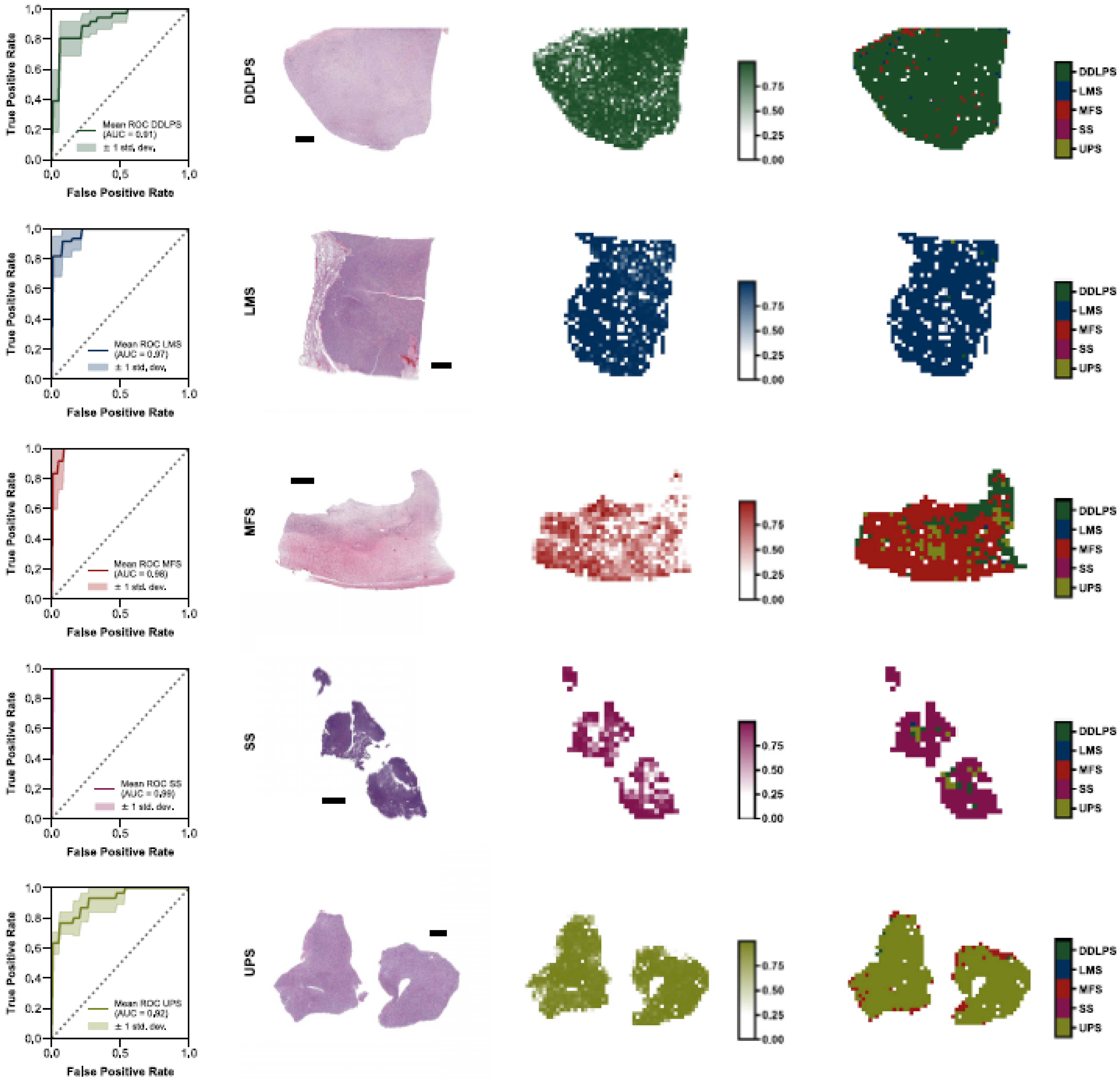
S. Foersch[1*], M. Eckstein[2], D.-C. Wagner[1], F. Gach[1], A.-C. Woerl[1,3], J. Geiger[1,3], C. Glasner[1,3], S. Schelbert[1], S. Schulz[1], S. Porubsky[1], A. Kreft[1], A. Hartmann[2], A. Agaimy[2] & W. Roth[1]

Volume 32 ■ Issue 9 ■ 2021

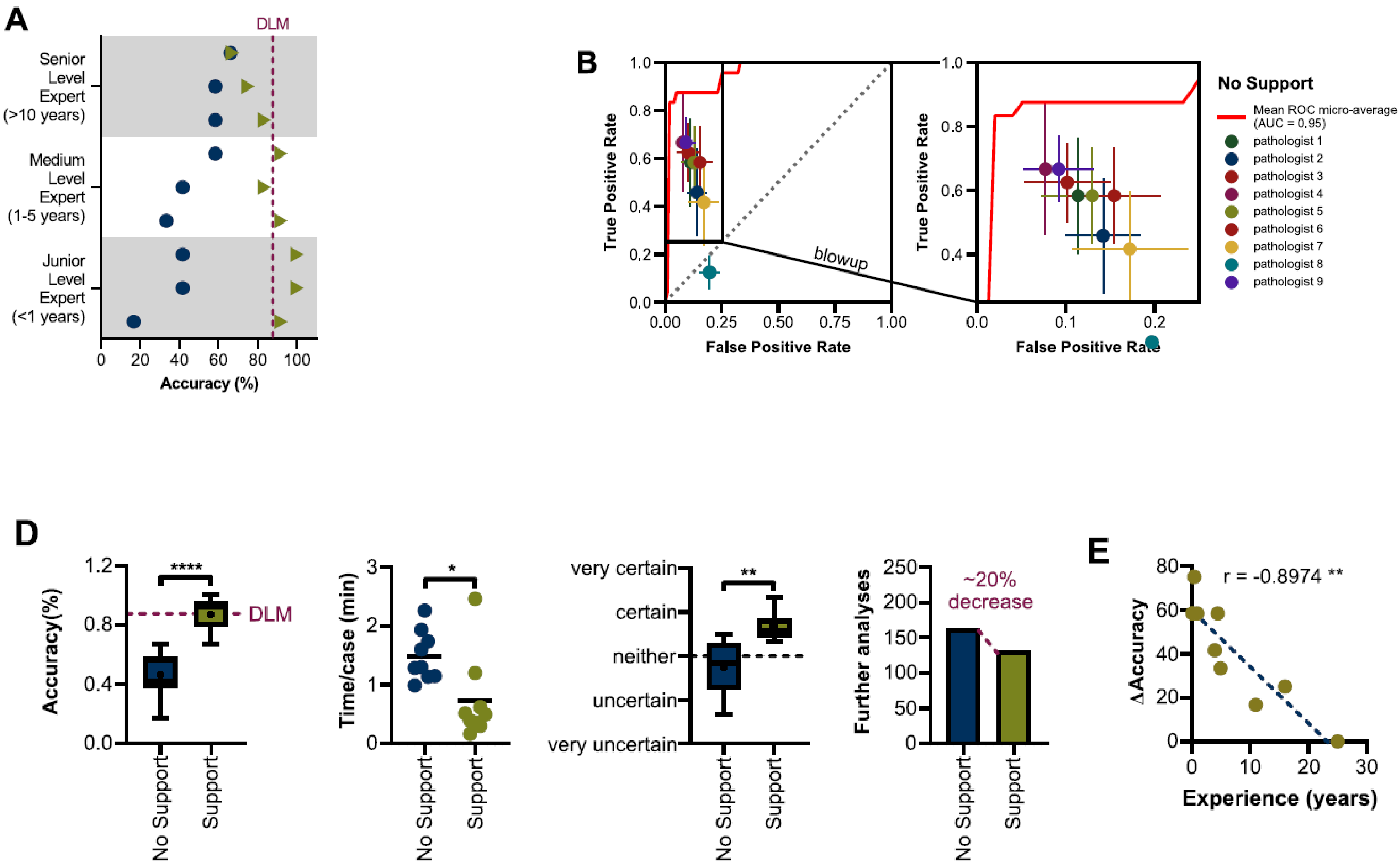
ANNALS OF ONCOLOGY



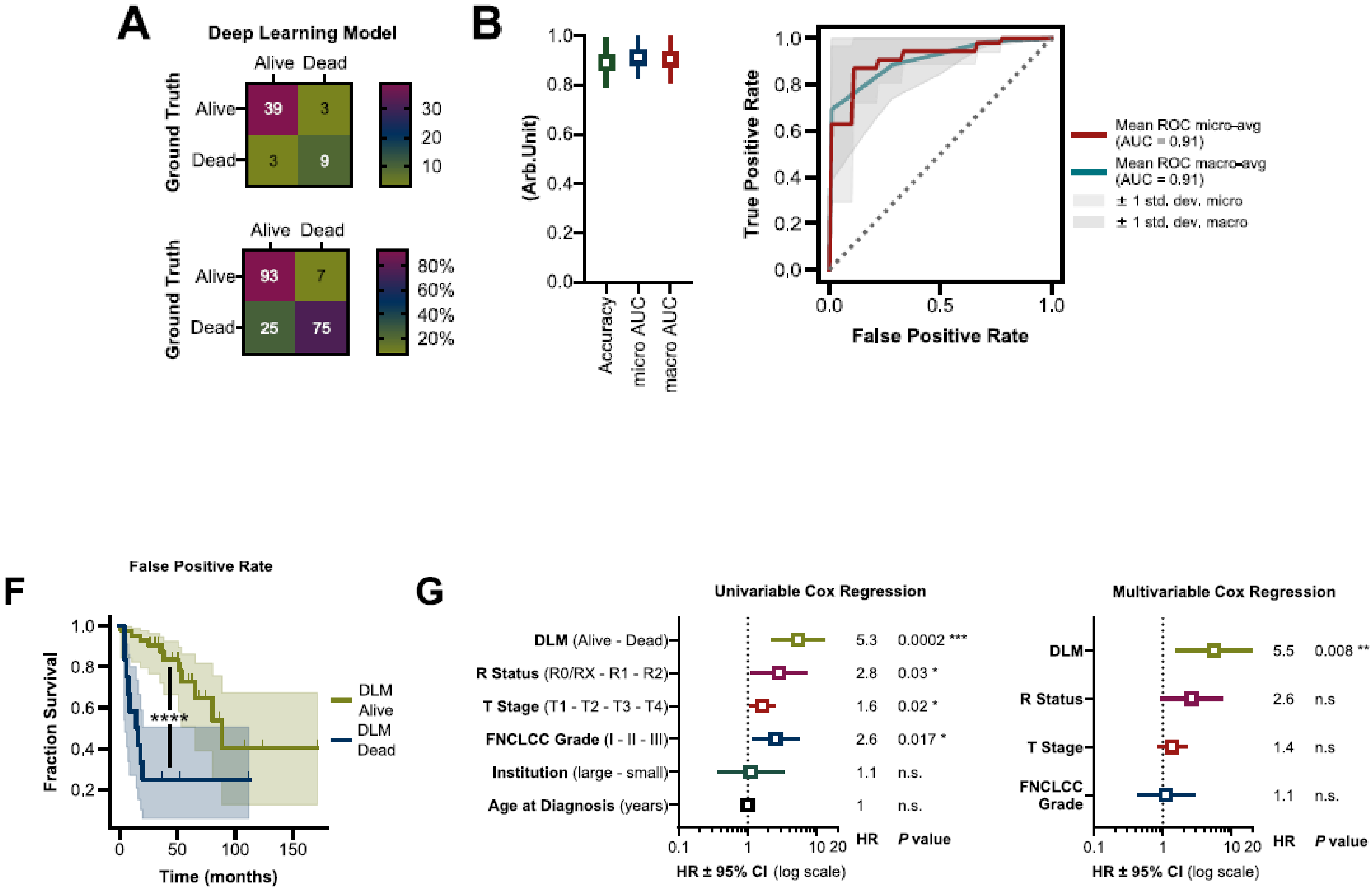

A
Deep Learning Model
DDLPS
LMS
MFS
SS
UPS
Ground Truth
26 2 2 2 4
1 50 2 2 5
0 0 10 0 2
0 0 0 6 0
3 0 4 0 23
50
40
30
20
10
0
Deep Learning Model
72 6 6 6 11
2 83 3 3 8
0 0 83 0 17
0 0 0 100 0
10 0 13 0 77
100%
80%
60%
40%
20%
0%
B
(Arb. Unit)
1.00
0.95
0.90
0.85
0.80
0.75
0.70
Accuracy
micro AUC
macro AUC
True Positive Rate
False Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.5
1.0
Mean ROC micro-avg (AUC = 0.95)
Mean ROC macro-avg (AUC = 0.97)
± 1 std. dev. micro
± 1 std. dev. macro



True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
0.5
Mean ROC DDLPS (AUC = 0.91)
± 1 std. dev.
Mean ROC LMS (AUC = 0.97)
± 1 std. dev.
Mean ROC MFS (AUC = 0.98)
± 1 std. dev.
Mean ROC SS (AUC = 0.99)
± 1 std. dev.
Mean ROC UPS (AUC = 0.92)
± 1 std. dev.
DDLPS
LMS
MFS
SS
UPS
0.00
0.25
0.50
0.75
DDLPS
LMS
MFS
SS
UPS



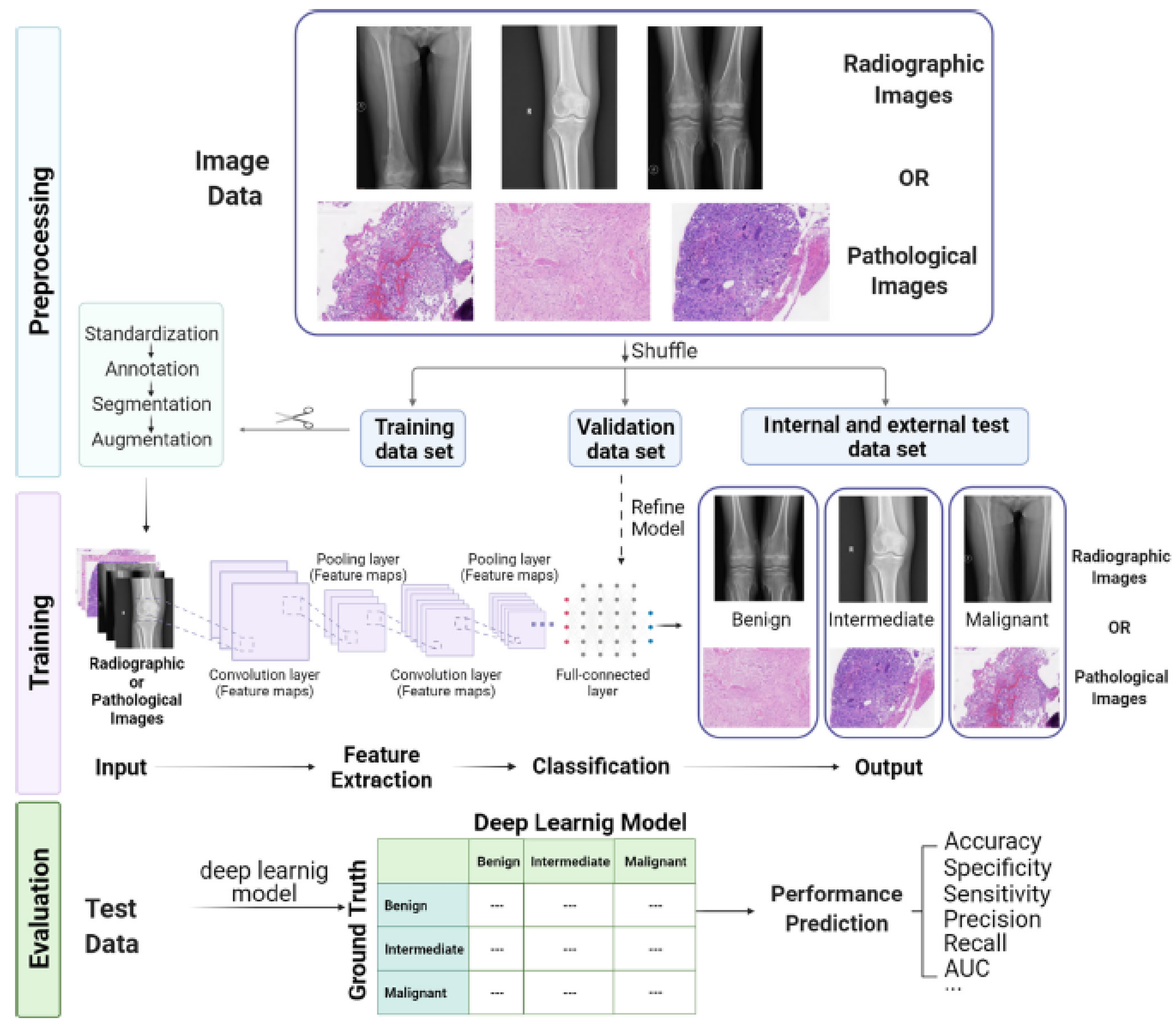
A
DLM
Senior Level Expert (>10 years)
Medium Level Expert (1-5 years)
Junior Level Expert (<1 years)
Accuracy (%)
B
True Positive Rate
False Positive Rate
blowup
No Support
Mean ROC micro-average (AUC = 0.95)
pathologist 1
pathologist 2
pathologist 3
pathologist 4
pathologist 5
pathologist 6
pathologist 7
pathologist 8
pathologist 9
D
Accuracy(%)
DLM
No Support
Support
Time/case (min)
very certain
certain
neither
uncertain
very uncertain
Further analyses
~20% decrease
E
r = -0.8974 **
ΔAccuracy
Experience (years)


A
Deep Learning Model
Alive Dead
Ground Truth
Alive
Dead
39 3
3 9
30
20
10
Alive Dead
Ground Truth
Alive
Dead
93 7
25 75
80%
60%
40%
20%
B
(Arb.Unit)
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
micro AUC
macro AUC
True Positive Rate
False Positive Rate
0.0 0.5 1.0
Mean ROC micro-avg (AUC = 0.91)
Mean ROC macro-avg (AUC = 0.91)
± 1 std. dev. micro
± 1 std. dev. macro
False Positive Rate
F
Fraction Survival
Time (months)
0 50 100 150
****
DLM Alive
DLM Dead
G
Univariable Cox Regression
DLM (Alive - Dead)
R Status (R0/RX - R1 - R2)
T Stage (T1 - T2 - T3 - T4)
FNCLCC Grade (I - II - III)
Institution (large - small)
Age at Diagnosis (years)
5.3 0.0002 ***
2.8 0.03 *
1.6 0.02 *
2.6 0.017 *
1.1 n.s.
1 n.s.
0.1 1 10 20
HR
P value
HR ± 95% CI (log scale)
Multivariable Cox Regression
DLM
R Status
T Stage
FNCLCC Grade
5.5 0.008 **
2.6 n.s
1.4 n.s
1.1 n.s.
0.1 1 10 20
HR
P value
HR ± 95% CI (log scale)

Emerging Applications of Deep Learning in Bone Tumors: Current Advances and Challenges

*Xiaowen Zhou[1,2], Hua Wang[2], Chengyao Feng[1,3], Ruilin Xu[1,3], Yu He[4], Lan Li[5] and Chao Tu[1,3]**

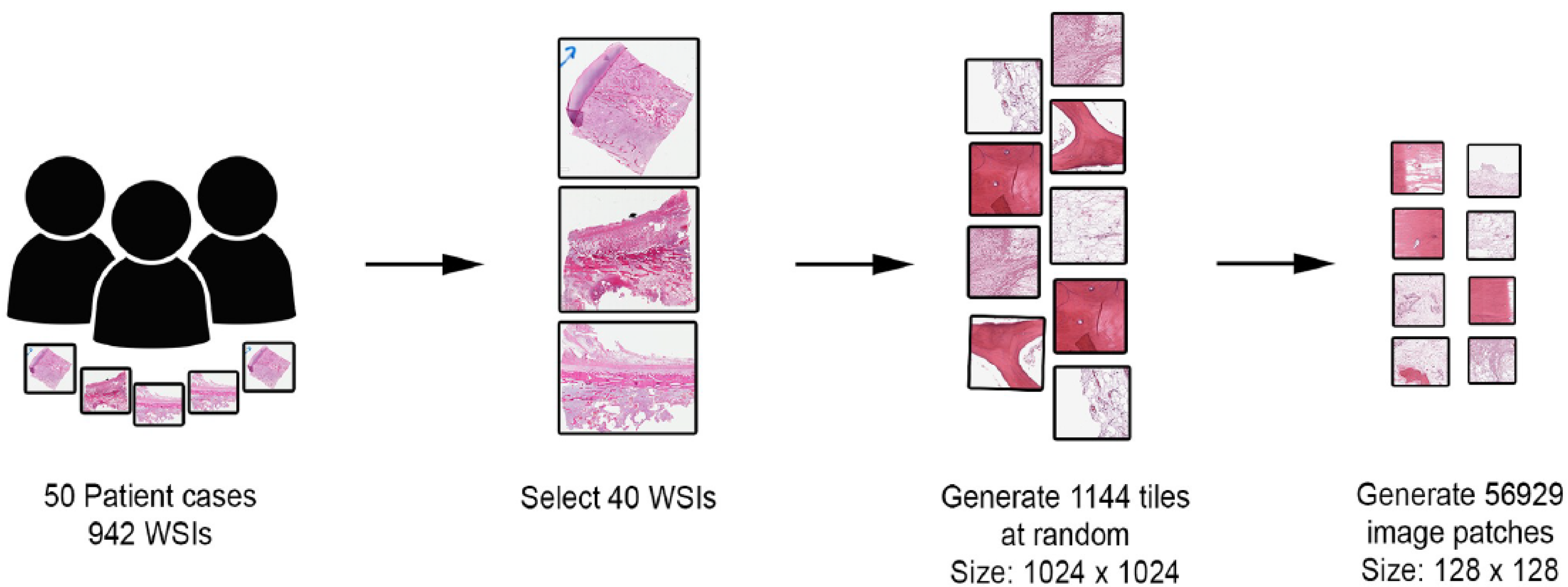
July 2022 | Volume 12 | Article 908873



Viable and necrotic tumor assessment from whole slide images of osteosarcoma using machine-learning and deep-learning models

Harish Babu Arunachalam[1☯], Rashika Mishra[1☯], Ovidiu Daescu[1*], Kevin Cederberg[2,3], Dinesh Rakheja[2,3], Anita Sengupta[2,3], David Leonard[3], Rami Hallac[3], Patrick Leavey[2,3]

PLOS ONE | https://doi.org/10.1371/journal.pone.0210706 April 17, 2019



Email: linosk@mskcc.org

@KonstantinosLin

https://www.youtube.com/channel/UCOgv4fITX9qVNu1Waf2ulRQ

Thank you

Memorial Sloan Kettering
Cancer Center