Deep Learning Model for Metastasis Risk in Colon Cancer Patients based on Binary Tumor Images

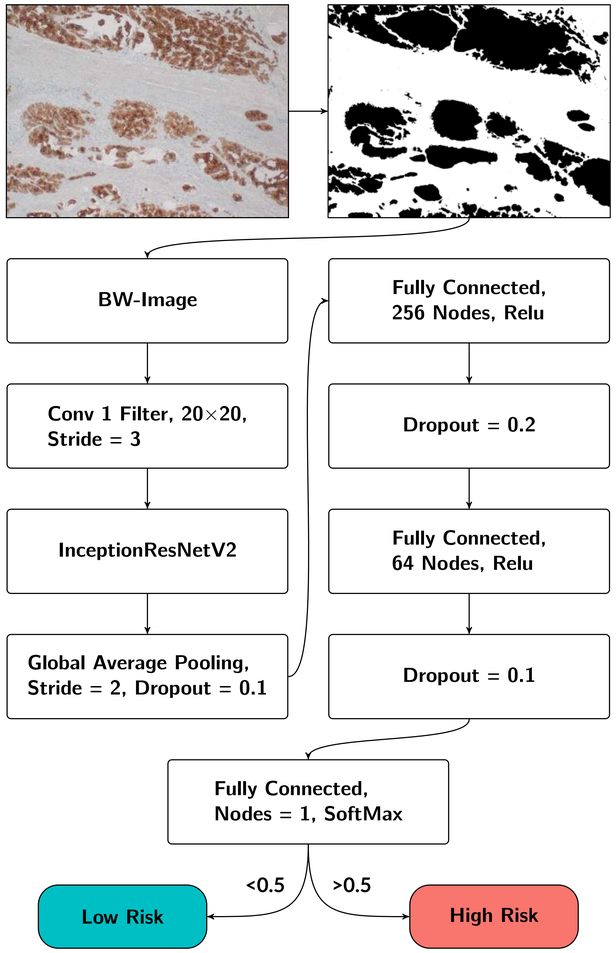
S. Schiele[1], T.T. Arndt[1,2], B. Martin[2], S. Miller[2], B. Märkl[2], G. Müller[1]
[1]Institute of Mathematics, Augsburg University, Germany
[2]General Pathology and Molecular Diagnostics, Medical Faculty, Augsburg, Germany

Introduction

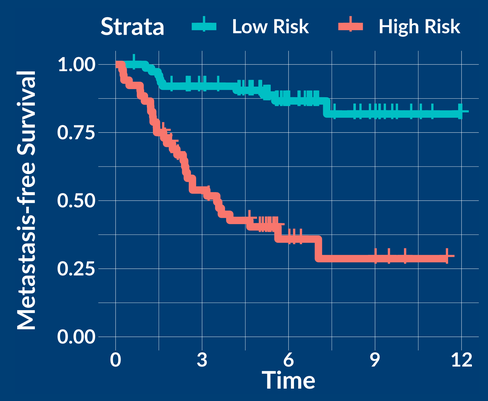
The goal of this study was to develop a CNN-based model for stratification of colon-cancer patients in two risk-groups regarding the occurrence of distant metastasis. To achieve this a model based on InceptionResNetV2 was trained on binary, histological images.

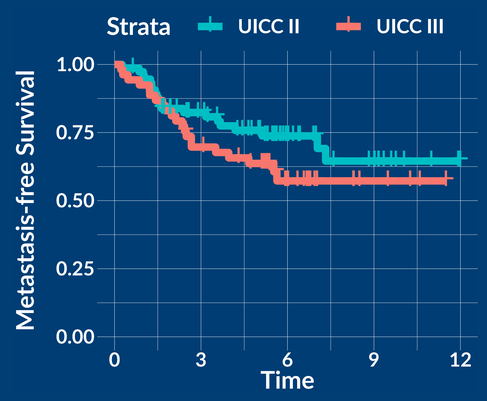
Material and Methods

We considered 291 patients with pT3- and pT4 adenocarcinoma, of no special histological type and without metastasis at diagnosis. For every patient a representative tumor section was chosen, stained with Cytokeratin AE1/AE3, and transformed to binary images using ImageJ. To avoid overfitting, the images were augmented during training and dropout layers were added to the model. The final output was used to predict a probability of distant metastasis and stratify the patients into a low and a high risk group.



Tumor architecture seems to be a strong predictor (log-rank test: p<0.001) for the risk of distant metastasis and outperforms UICC staging.





Conclusion

Deep Learning can stratify colon cancer patients well based only on black and white images of the tumor architecture. Further investigation to find interesting features for the decision of the model.

Universität Augsburg University

CCCA UNIVERSITÄTSKLINIKUM AUGSBURG Comprehensive Cancer Center Augsburg

Results

Patient characteristics in the validation collective (n = 128)

Variable	Val. Set	Low Risk	High Risk	p value
Sociodemographic characteristics				
Age, mean (SD), y	69 (12)	69 (12)	69 (12)	0.753
Sex, n (%)				0.764
Female	50 (39)	31 (41)	19 (37)	
Male	78 (61)	45 (59)	33 (64)	
Followup duration, median years	5.8	5.9	5.5	0.780
Clinicopathological characteristics				
Tumor proportion, mean (SD)	0.36 (0.18)	0.38 (0.17)	0.33 (0.20)	0.143
Tumor proportion, n (%)				0.002
Low	21 (16)	6 (8)	15 (29)	
Medium	80 (63)	56 (74)	24 (46)	
High	27 (21)	14 (18)	13 (25)	
Tumor staging, n (%)				0.056
pT3	109 (85)	69 (91)	40 (77)	
pT4	19 (15)	7 (9)	12 (23)	
Nodal status, n (%)				0.278
Negative	75 (59)	48 (63)	27 (52)	
Positive	53 (41)	28 (37)	25 (48)	
Lymphovascular invasion, n (%)				1.000
Negative	104 (81)	62 (82)	42 (81)	
Positive	24 (19)	14 (18)	10 (19)	
Tumor budding, n (%)				0.328
Bd 1	103 (80)	64 (84)	39 (75)	
Bd 2	15 (12)	8 (11)	7 (13)	
Bd 3	10 (8)	4 (5)	6 (12)	
Location of tumor, n (%)				0.614
Right	76 (59)	47 (62)	29 (56)	
Left	52 (41)	29 (38)	23 (44)	
Microsatellite status, n (%)				0.896
Stable	115 (90)	69 (91)	46 (88)	
Instable	13 (10)	7 (9)	6 (12)	
UICC, n (%)				0.278
II	75 (59)	48 (63)	27 (52)	
III	53 (41)	28 (37)	25 (48)	
Distant Metastasis, n (%)				<0.001
Yes	41 (32)	10 (13)	31 (60)	
No	87 (68)	66 (87)	21 (40)	

- **Good discrimination** of patients according to occurrence of metastasis (AUC: 0.842 95%-CI: 0.774-0.911)
- **Risk-classification** was a **prognostic factor** in an adjusted multivariate Cox-regression
- Useful stratification **especially in UICC III subgroup**, where 80% of high-risk group developed metastasis, whereas in the low risk group only 4%.

Variable	HR (95% CI)	p-Value
Sociodemographic characteristics		
Age (continuous)	1.01 (0.98–1.04)	0.592
Sex (ref.: female)	1.2 (0.6–2.6)	0.626
Clinicopathological characteristics		
Predicted risk group (ref.: low)	5.4 (2.5–11.7)	<0.001
Tumor proportion (ref.: High)		
Medium	0.4 (0.2–1.0)	0.047
Low	0.7 (0.3–1.7)	0.410
Tumor staging (ref.: pT3)	2.6 (1.1–6.0)	0.029
Nodal status (ref.: negative)	0.9 (0.5–1.8)	0.838
Lymphovascular invasion (ref.: negative)	1.3 (0.6–3.2)	0.517
Tumor budding	1.6 (1.0–2.7)	0.069
Location of tumor (ref: right side)	1.5 (0.7–3.2)	0.245
Microsatellite status (ref.: Stable)	0.2 (0.0–1.2)	0.076