Ninth ISIC Skin Image Analysis Workshop @ MICCAI 2024

Segmentation Style Discovery: Application to Skin Lesion Images

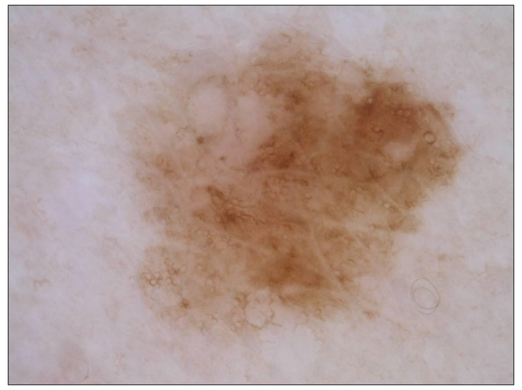


Kumar Abhishek[†] Jeremy Kawahara[‡] Ghassan Hamarneh[†]

[†] SFU Simon Fraser University

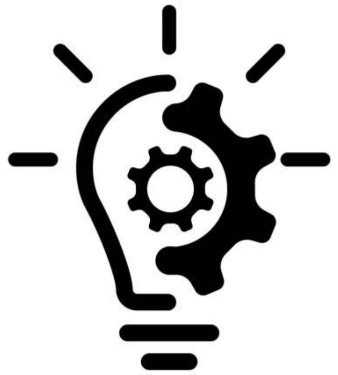
[‡] AIP LABS®

Variability in Medical Image Segmentation



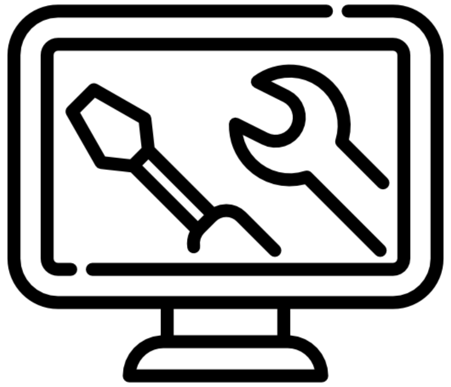
Ambiguous object boundaries

Annotators' personal preferences



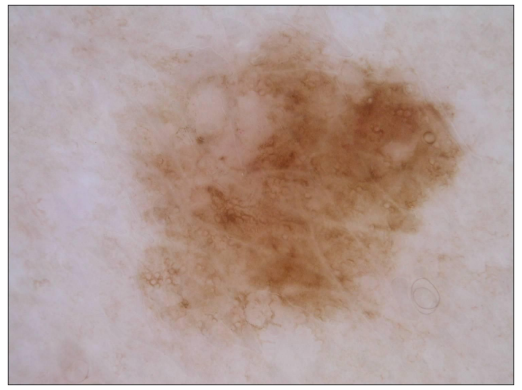
Annotators' skill levels

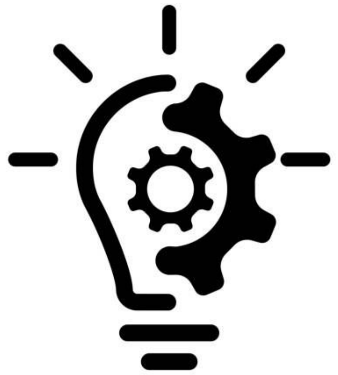
Segmentation criteria

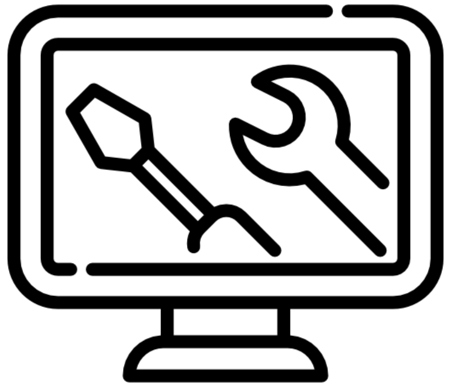


Segmentation tools

Variability in Medical Image Segmentation







Ambiguous object boundaries

Annotators' personal preferences

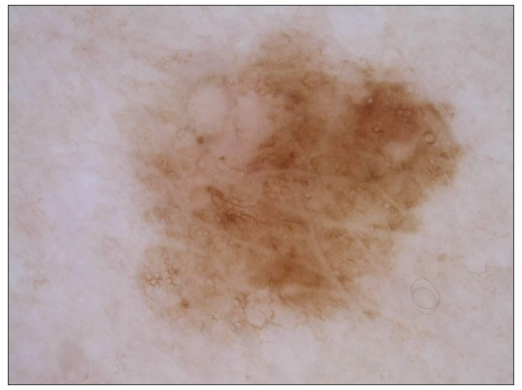
Annotators' skill levels

Segmentation criteria

Segmentation tools

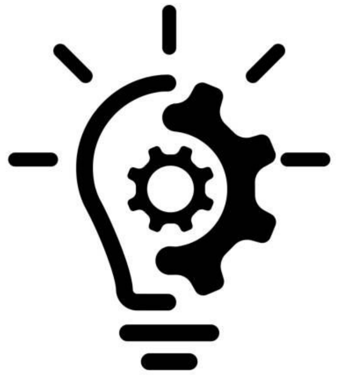
Latent factors

Variability in Medical Image Segmentation



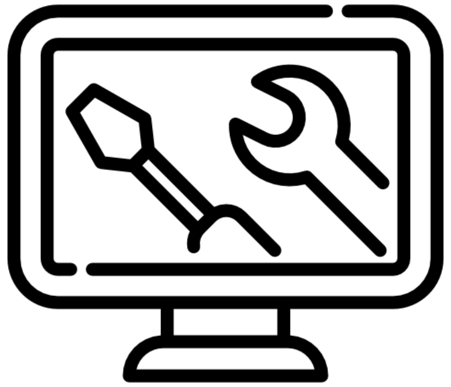
Ambiguous object boundaries

Annotators' personal preferences



Annotators' skill levels

Segmentation criteria



Segmentation tools

Latent factors

Different annotation segmentation preferences or "**styles**"

Methods for Learning from Multiple Annotations

$\mathbb{SS}\mathbf{eg}$ **methods** model and learn to predict a single “gold standard” segmentation.

[1] Silletti et al., 2009

Methods for Learning from Multiple Annotations

SSeg **methods** model and learn to predict a single “gold standard” segmentation.

***Abstract*— In a double blind evaluation of 60 digital dermatoscopic images by 4 “junior”, 4 “senior” and 4 “expert” dermatologists (dermatoscopy training respectively less than 1 year, between 1 and 5 years, and more than 5 years), a significant inter-operator variability was observed in melanocytic lesion border identification (with a disagreement of the order of $10-20\%$ of the area of the lesions). Expert dermatologists**

dard operative definition. For example, if even experienced dermatologists disagree on how to classify 5% of the area of an image, no automated system can be expected to classify “correctly” more than 95% of the area of that image.

[1] Silletti et al., 2009

Methods for Learning from Multiple Annotations

$\mathbb{SSeg}$ **methods** model and learn to predict a single “gold standard” segmentation.

> ***Abstract*— In a double blind evaluation of 60 digital dermatoscopic images by 4 “junior”, 4 “senior” and 4 “expert” dermatologists (dermatoscopy training respectively less than 1 year, between 1 and 5 years, and more than 5 years), a significant inter-operator variability was observed in melanocytic lesion border identification (with a disagreement of the order of $10 - 20\%$ of the area of the lesions). Expert dermatologists**

> dard operative definition. For example, if even experienced dermatologists disagree on how to classify 5% of the area of an image, no automated system can be expected to classify “correctly” more than 95% of the area of that image.

$\mathbb{MSeg}$ **methods** model and predict multiple segmentations to capture annotation variability.

[1] Silletti et al., 2009

Methods for Learning from Multiple Annotations

SSeg methods model and learn to predict a single "gold standard" segmentation.

> ***Abstract*—In a double blind evaluation of 60 digital dermatoscopic images by 4 "junior", 4 "senior" and 4 "expert" dermatologists (dermatoscopy training respectively less than 1 year, between 1 and 5 years, and more than 5 years), a significant inter-operator variability was observed in melanocytic lesion border identification (with a disagreement of the order of 10 – 20% of the area of the lesions). Expert dermatologists**

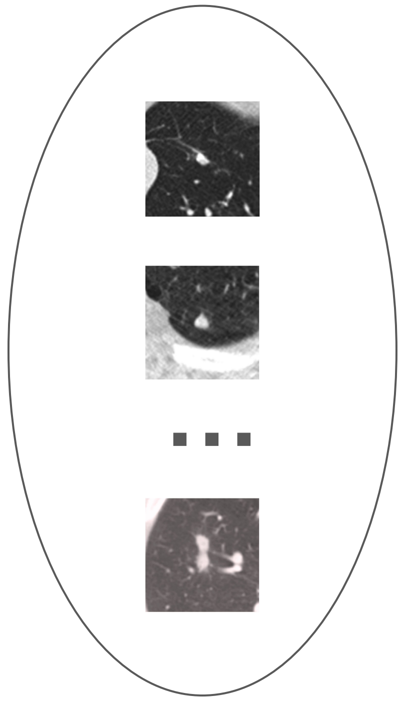
> dard operative definition. For example, if even experienced dermatologists disagree on how to classify 5% of the area of an image, no automated system can be expected to classify "correctly" more than 95% of the area of that image.

MSeg methods model and predict multiple segmentations to capture annotation variability.

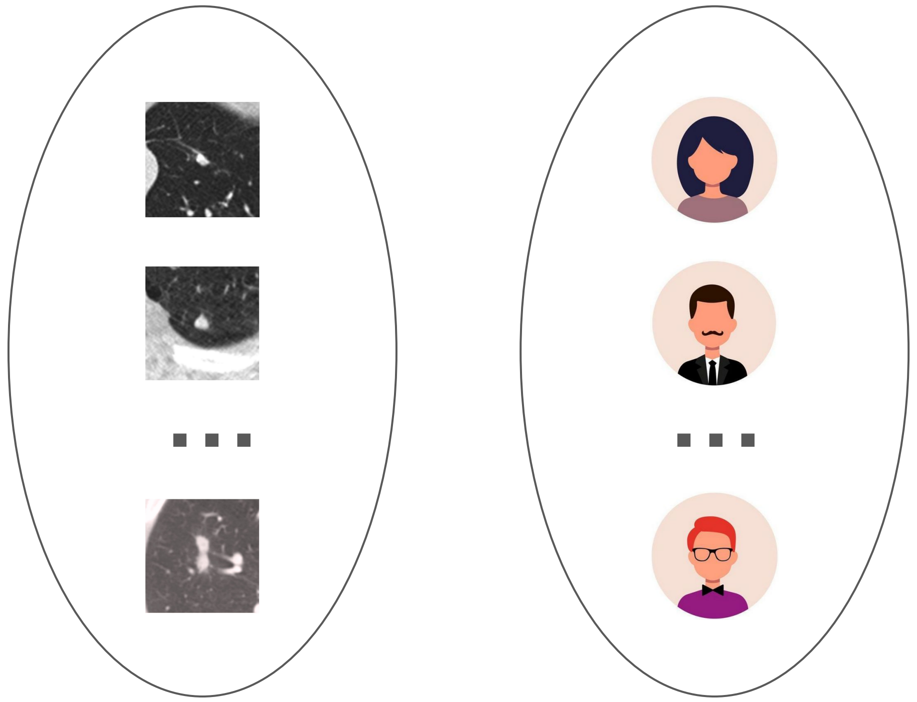
Dataset requirement:
multi-annotator segmentations containing image-mask pairs with annotator-segmentation correspondence.

[1] Silletti et al., 2009

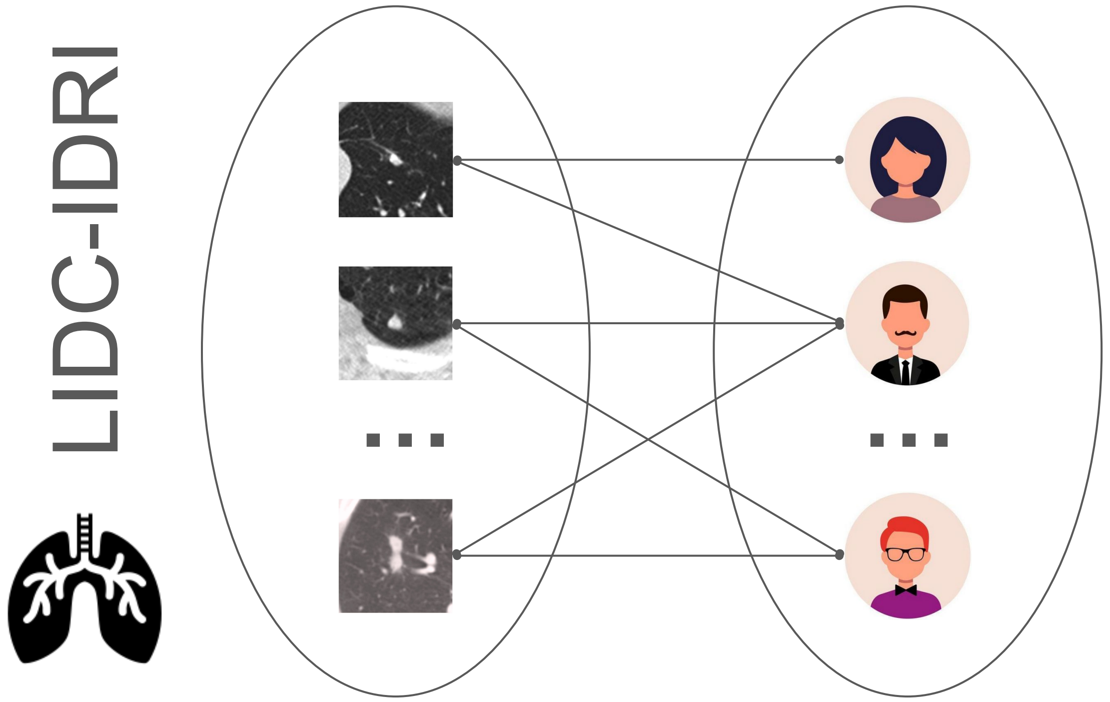
Multi-Annotator Medical Image Segmentation Datasets



Multi-Annotator Medical Image Segmentation Datasets



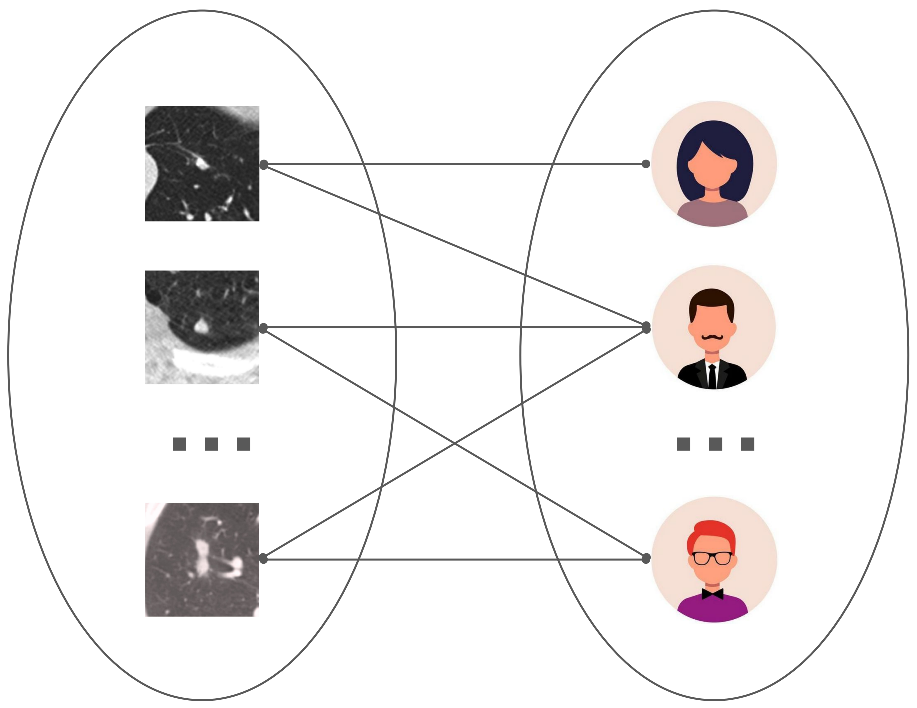
Multi-Annotator Medical Image Segmentation Datasets



Bipartite graph

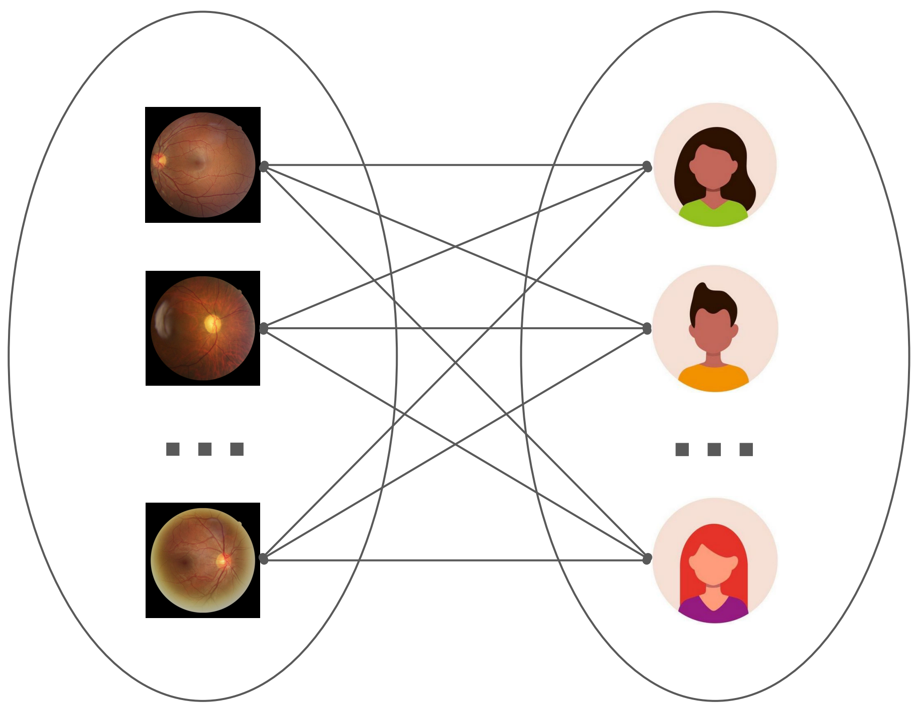
Multi-Annotator Medical Image Segmentation Datasets

LIDC-IDRI



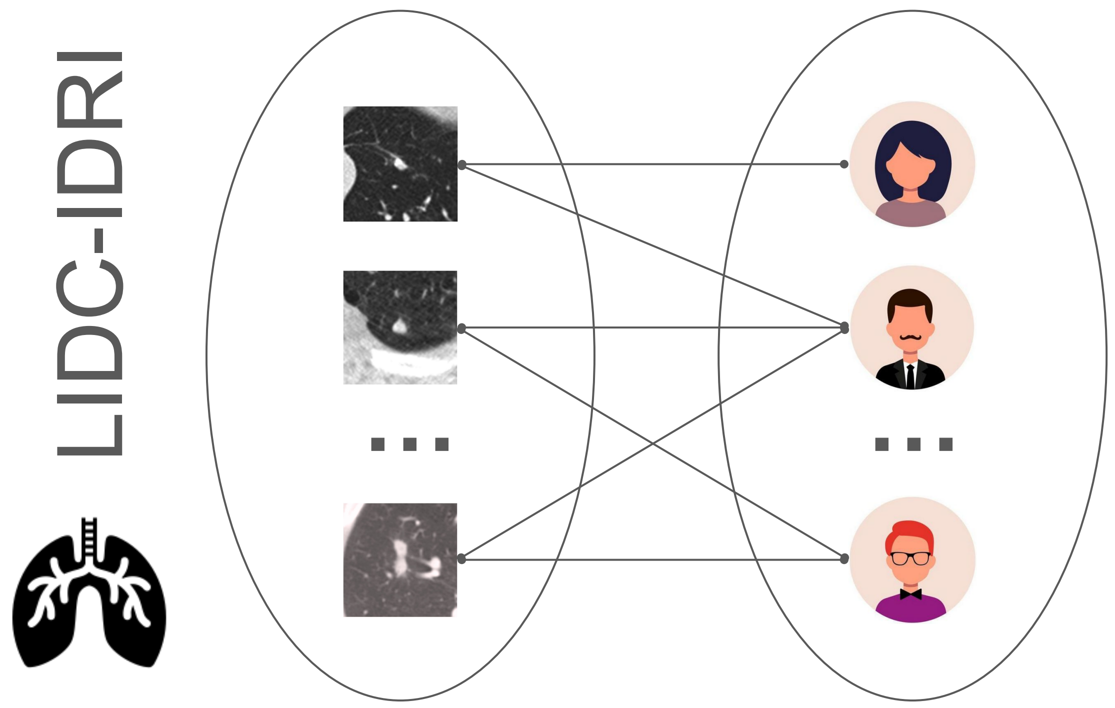
Bipartite graph

RIGA

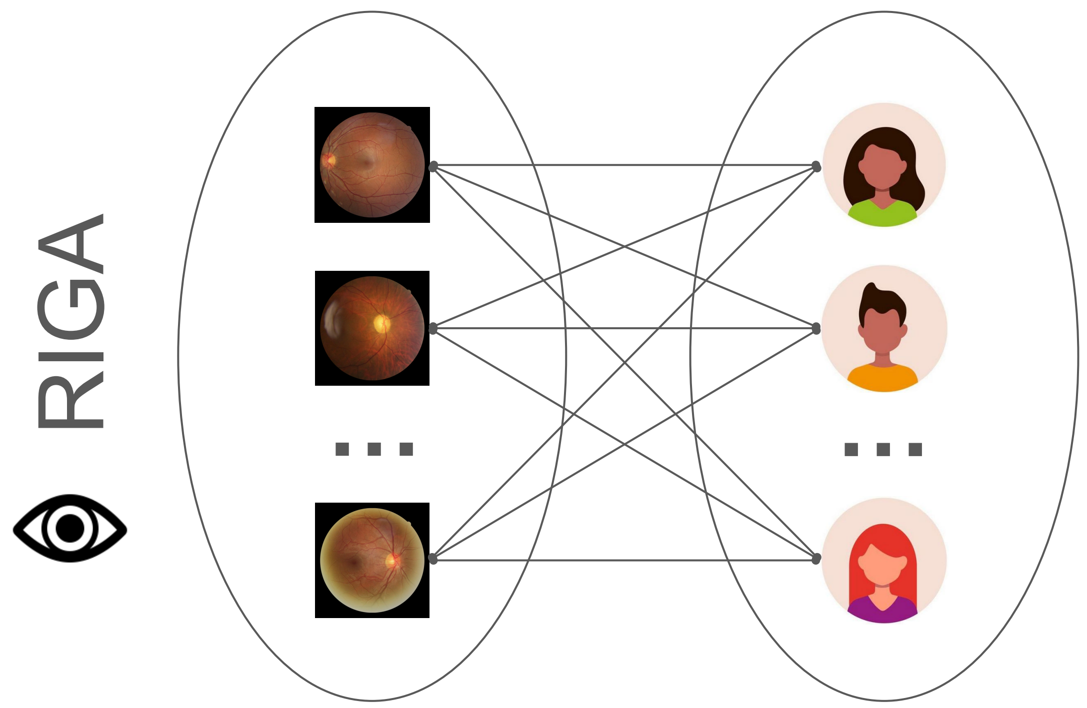


<u>Complete</u> bipartite graph

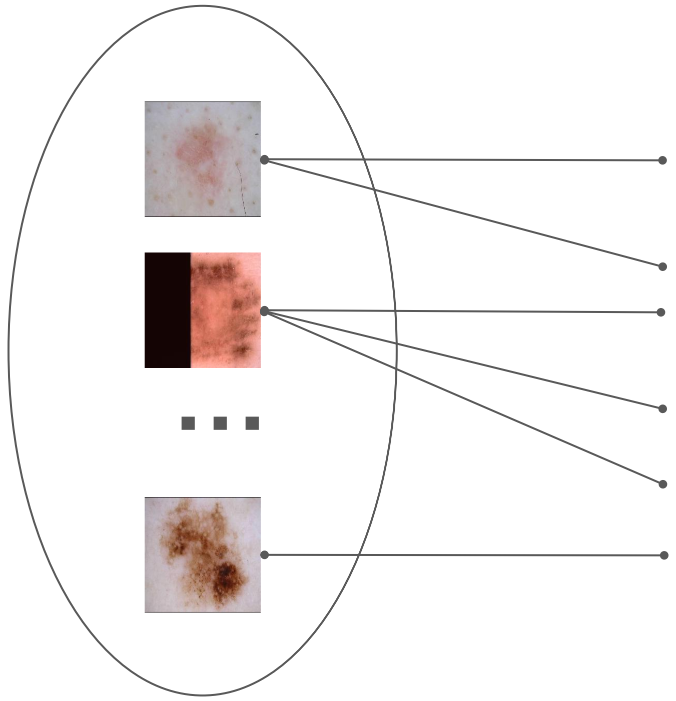
Multi-Annotator Medical Image Segmentation Datasets



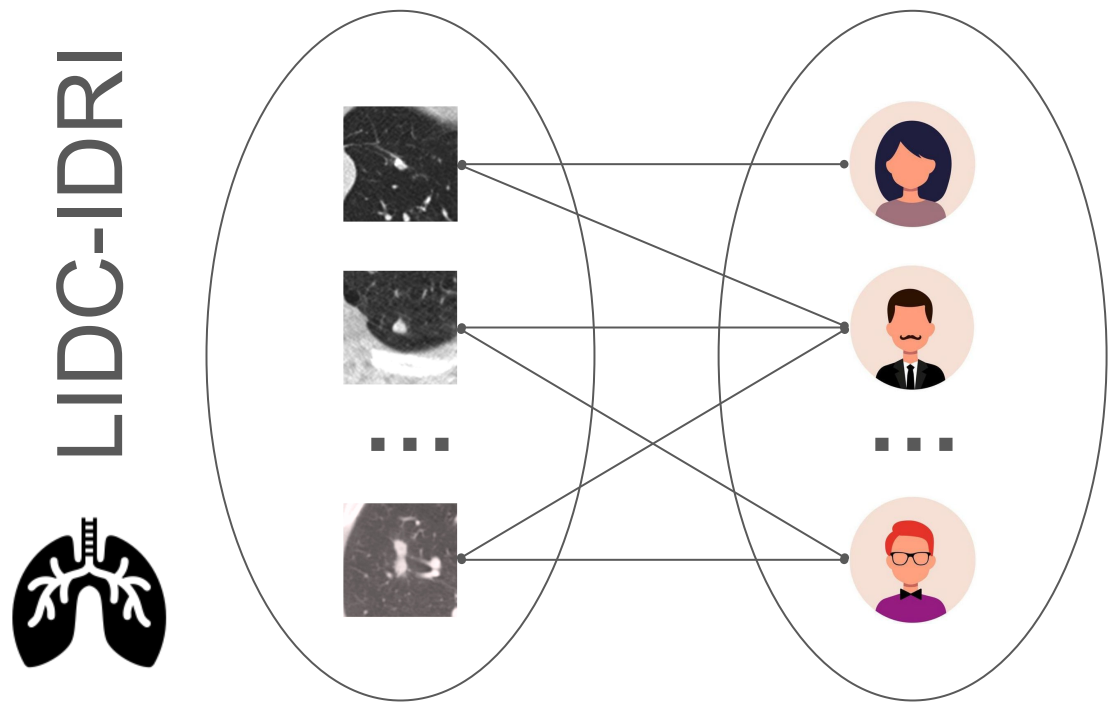
Bipartite graph



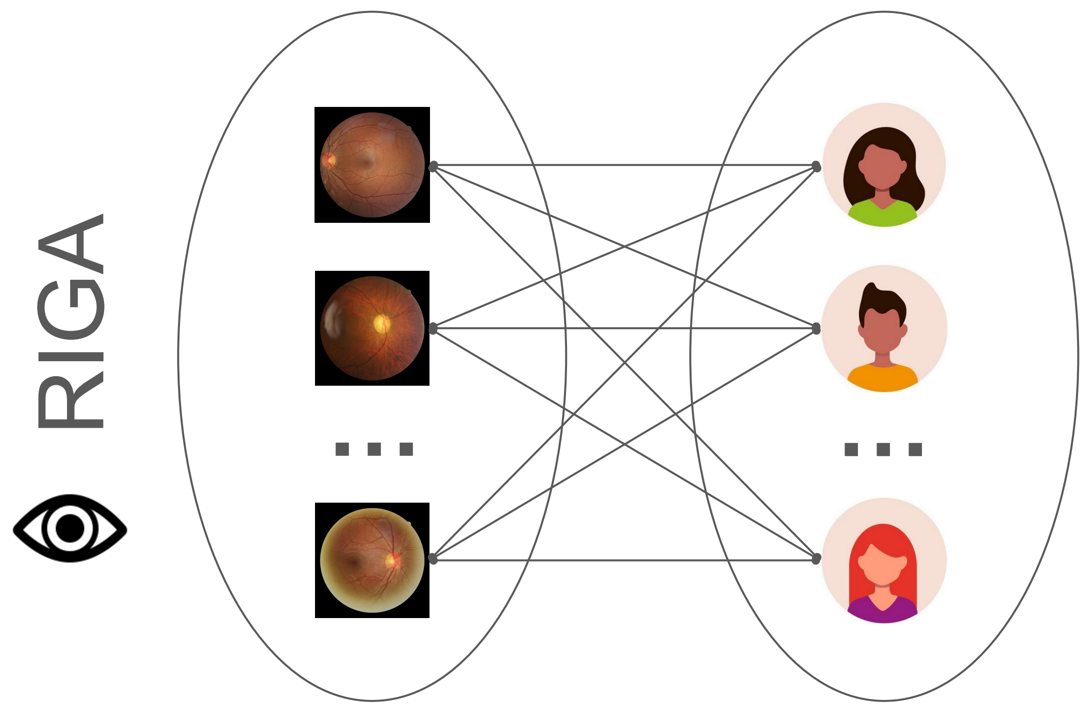
Complete bipartite graph



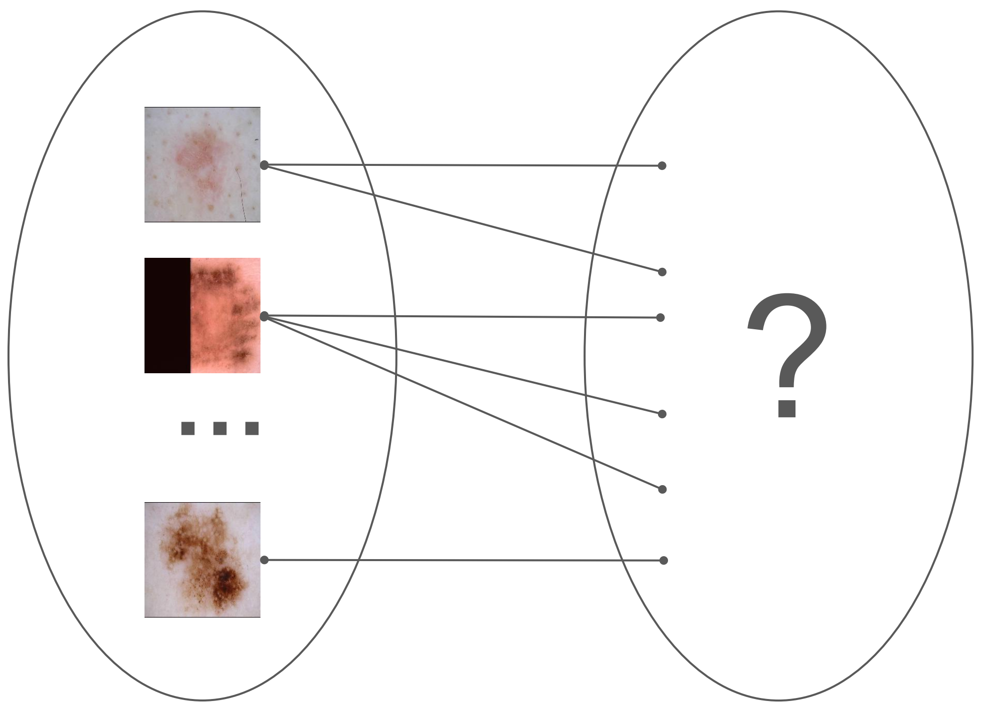
Multi-Annotator Medical Image Segmentation Datasets



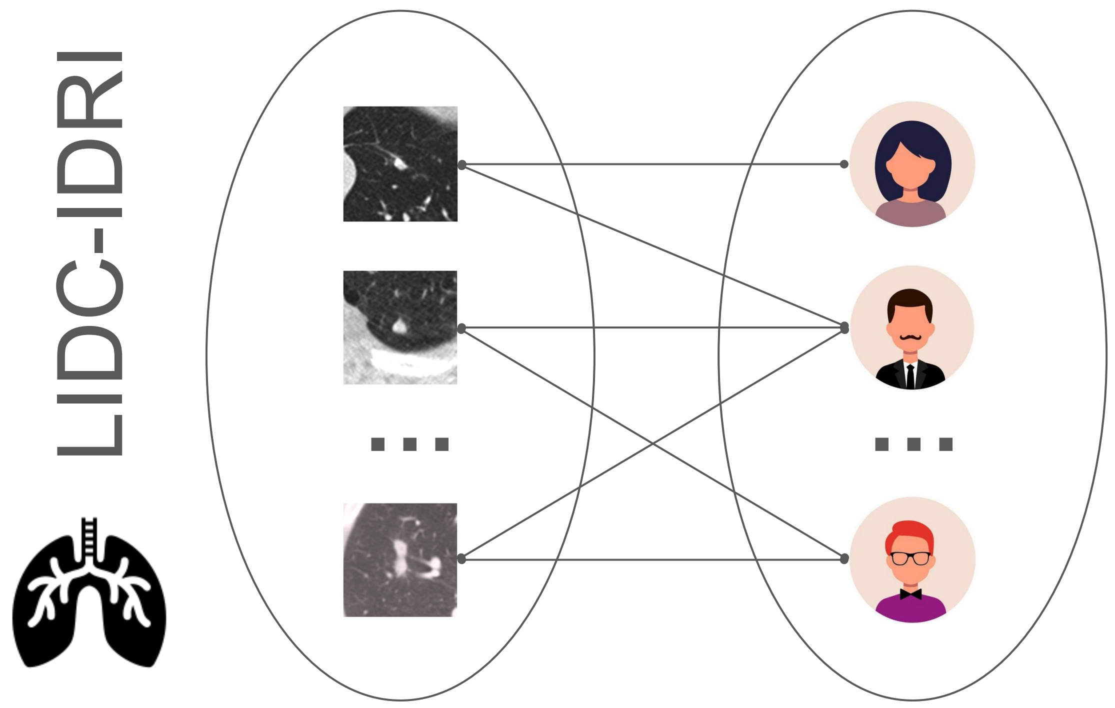
Bipartite graph



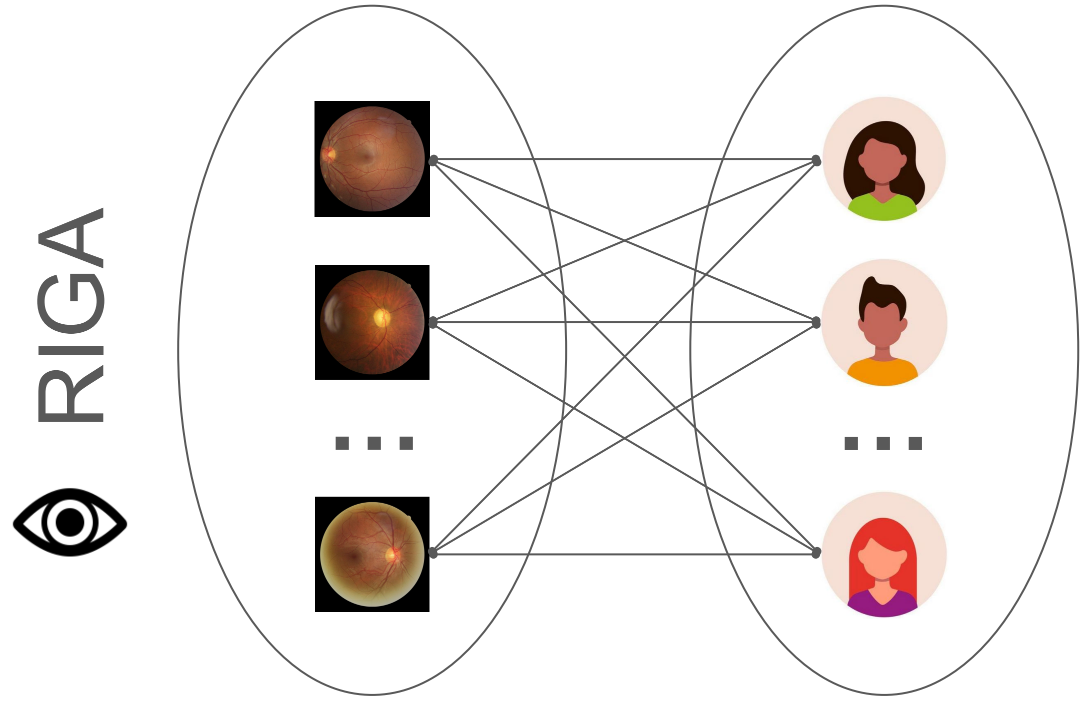
Complete bipartite graph



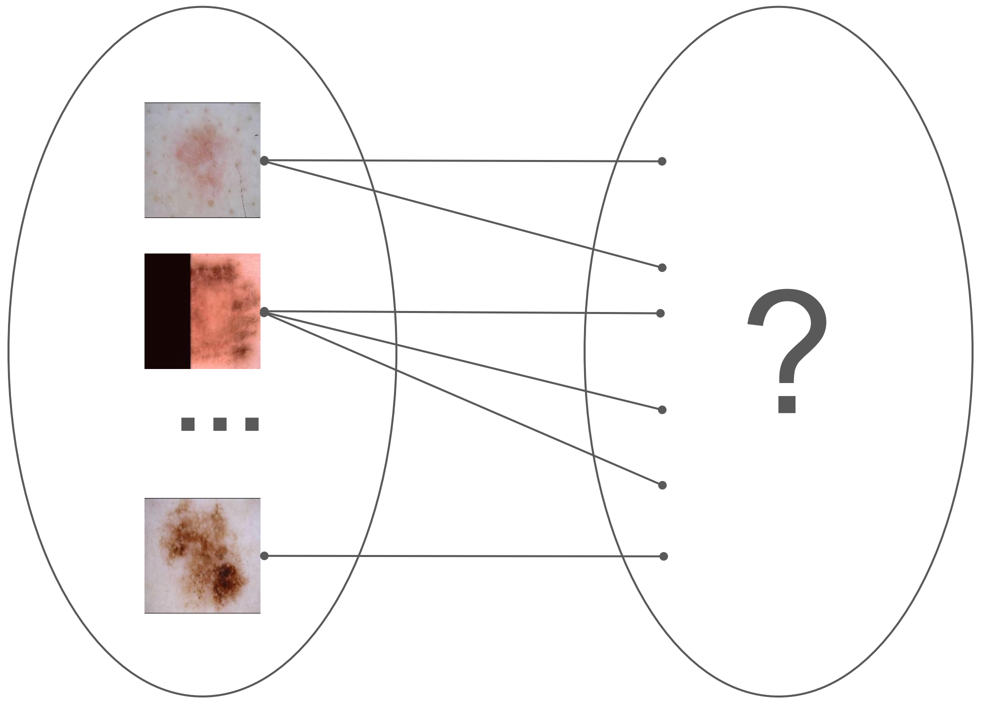
Multi-Annotator Medical Image Segmentation Datasets



Bipartite graph



<u>Complete</u> bipartite graph



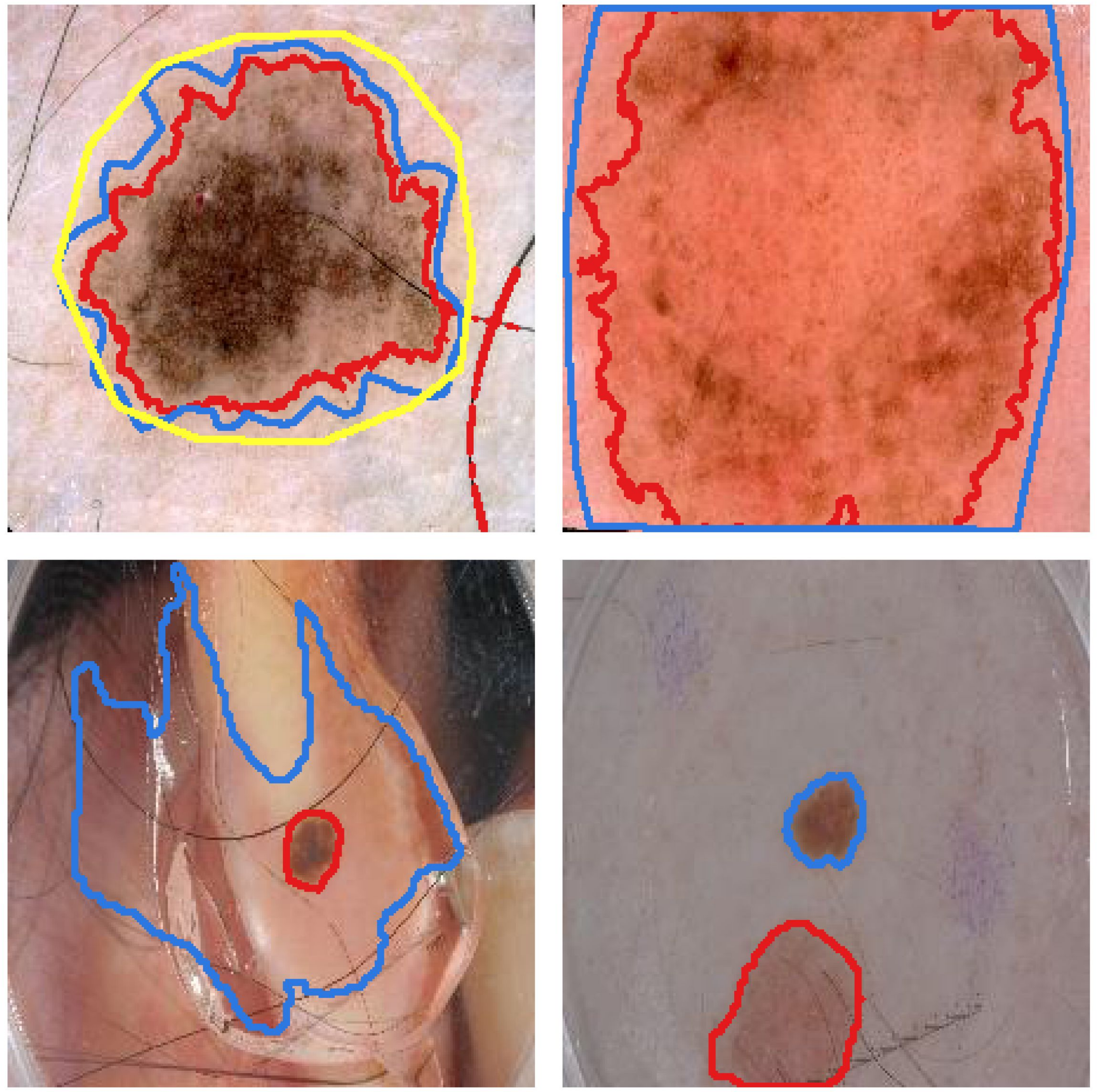
Latent factors unknown ⇒ difficult to define a segmentation “**style**”.

Segmentations in ISIC Archive and their Variability

2,261 images with more than 1 “ground truth” segmentation mask
⇒ 4,704 training image-mask pairs for skin lesion segmentation (**SLS**).

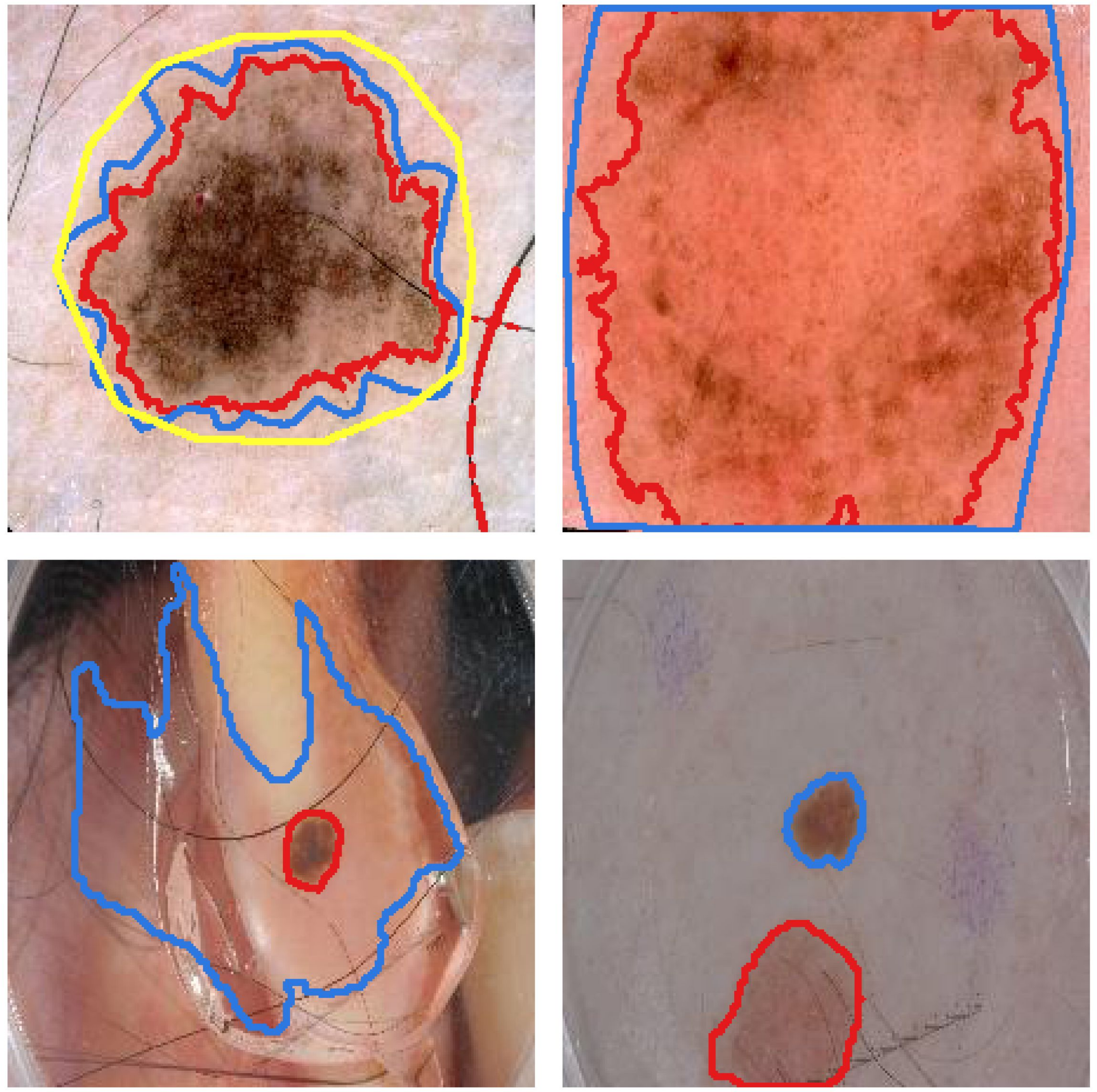
Segmentations in ISIC Archive and their Variability

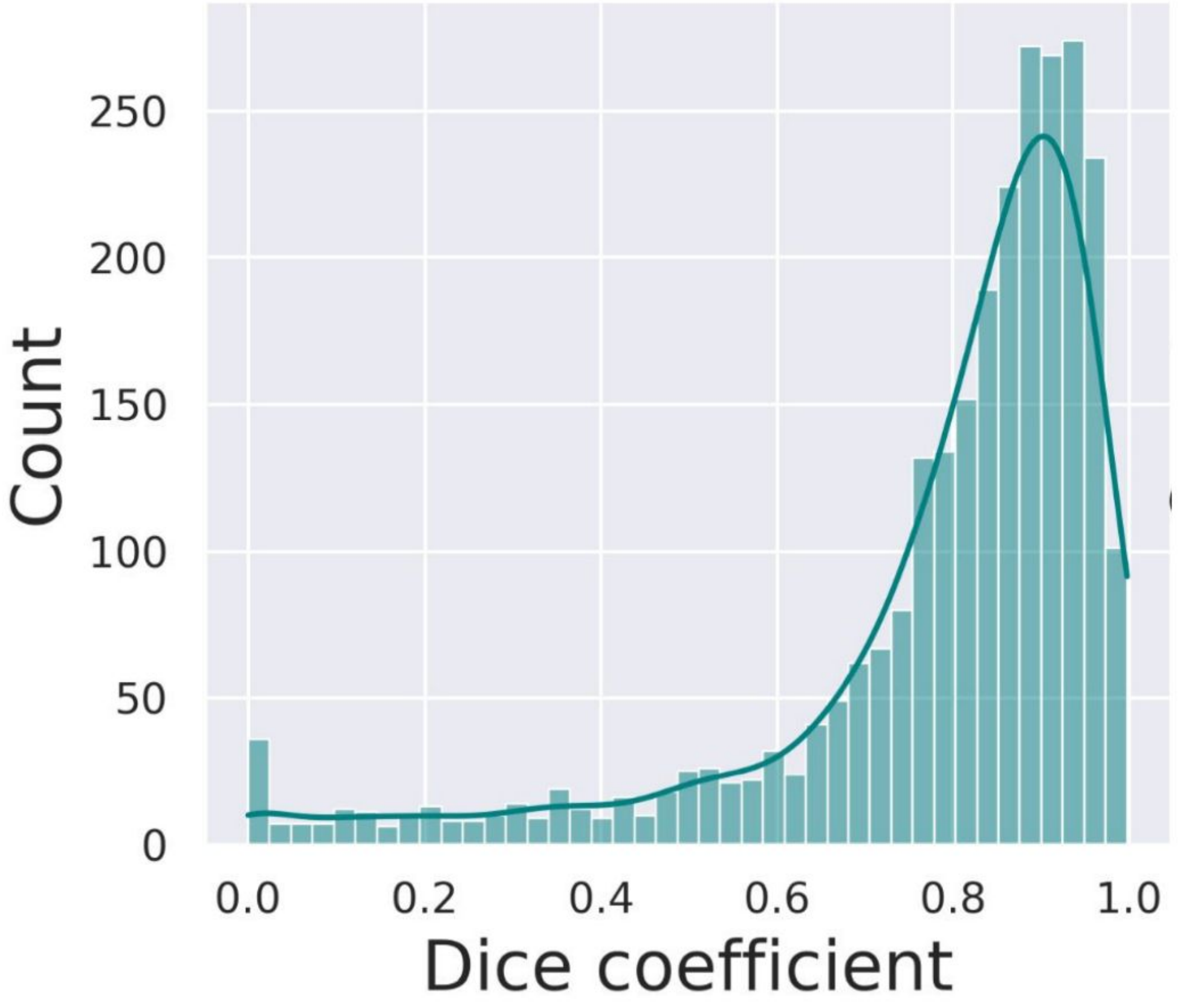
2,261 images with more than 1 “ground truth” segmentation mask
⇒ 4,704 training image-mask pairs for skin lesion segmentation (**SLS**).

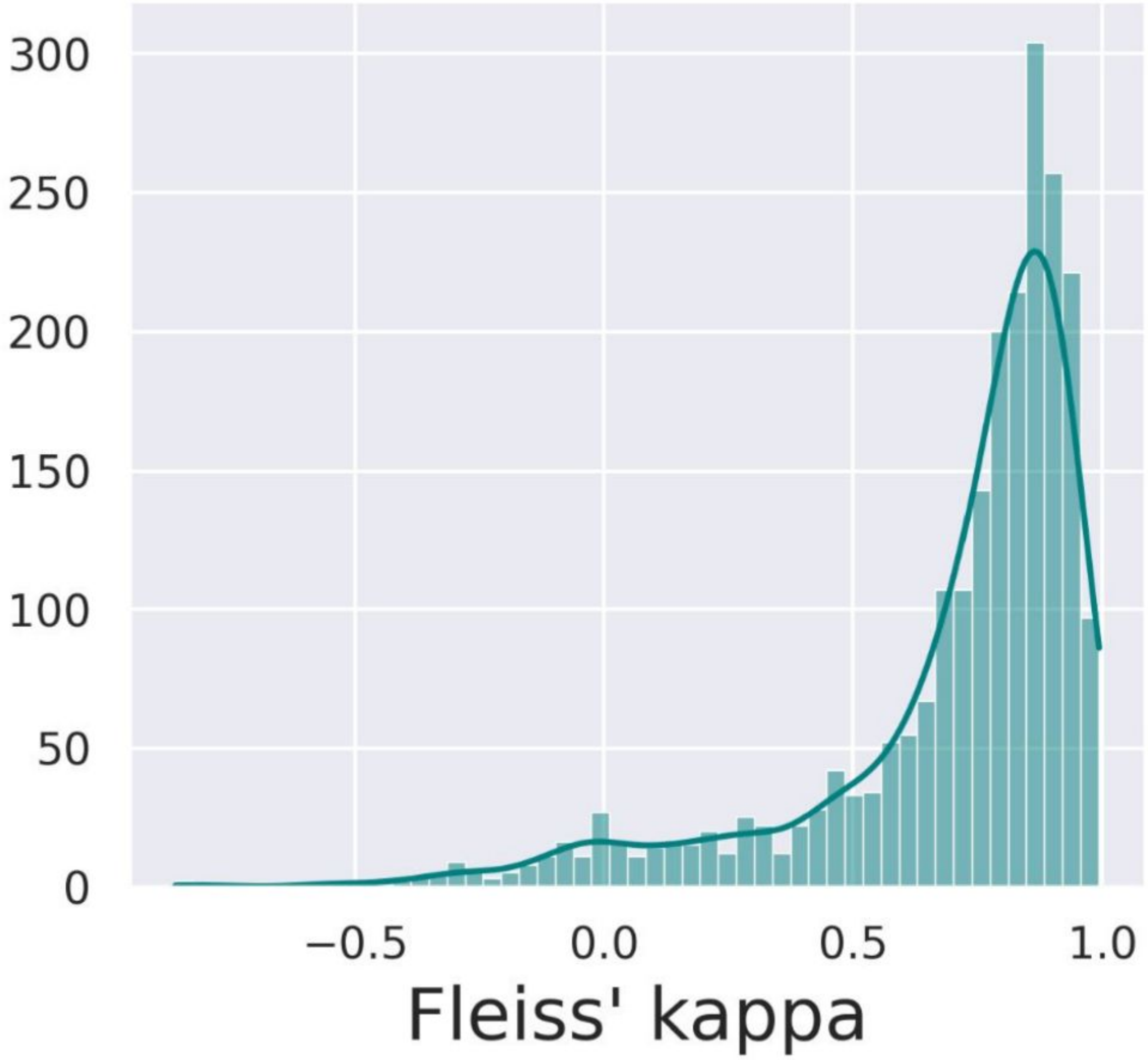


Segmentations in ISIC Archive and their Variability

2,261 images with more than 1 “ground truth” segmentation mask
⇒ 4,704 training image-mask pairs for skin lesion segmentation (**SLS**).

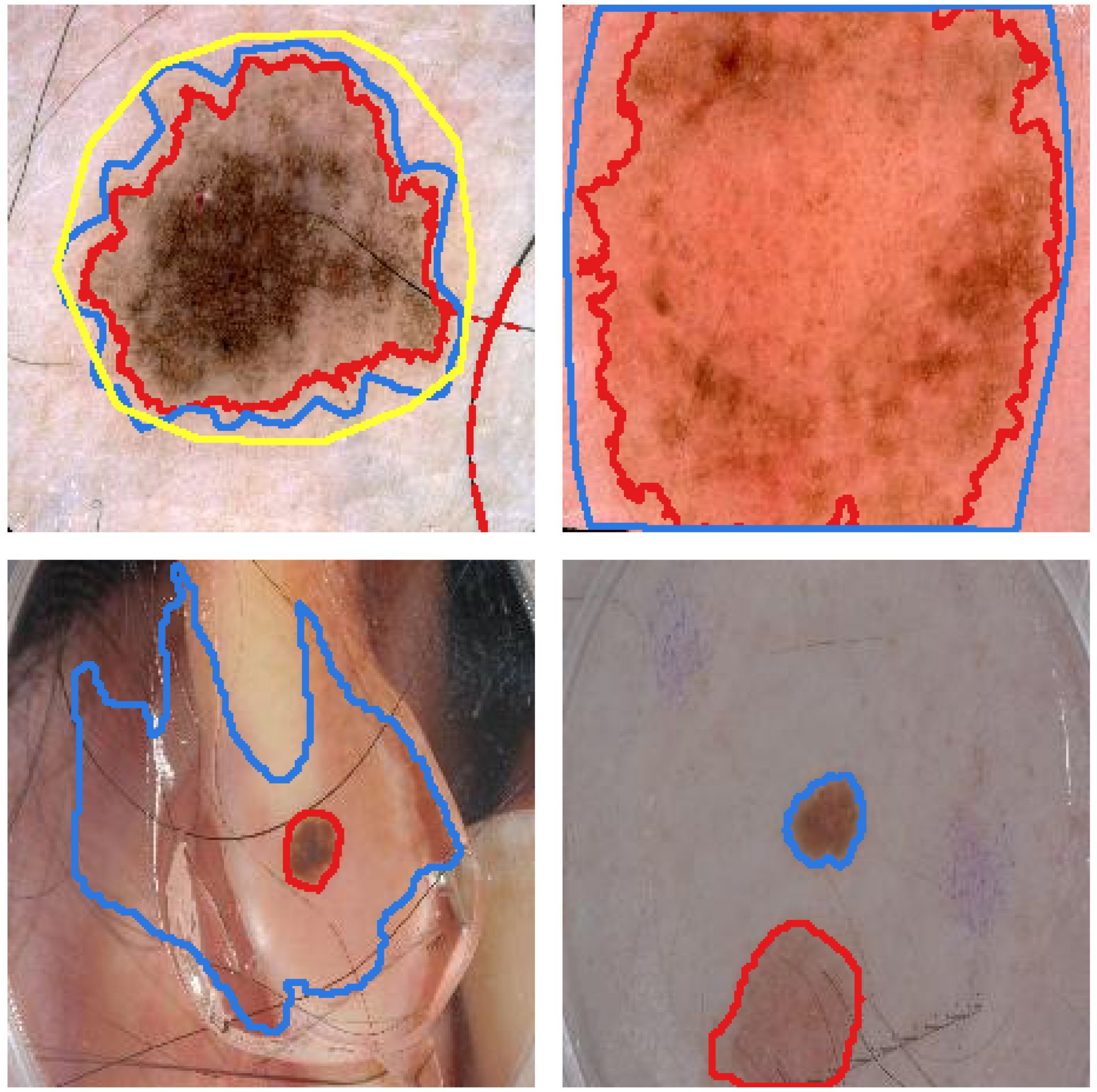


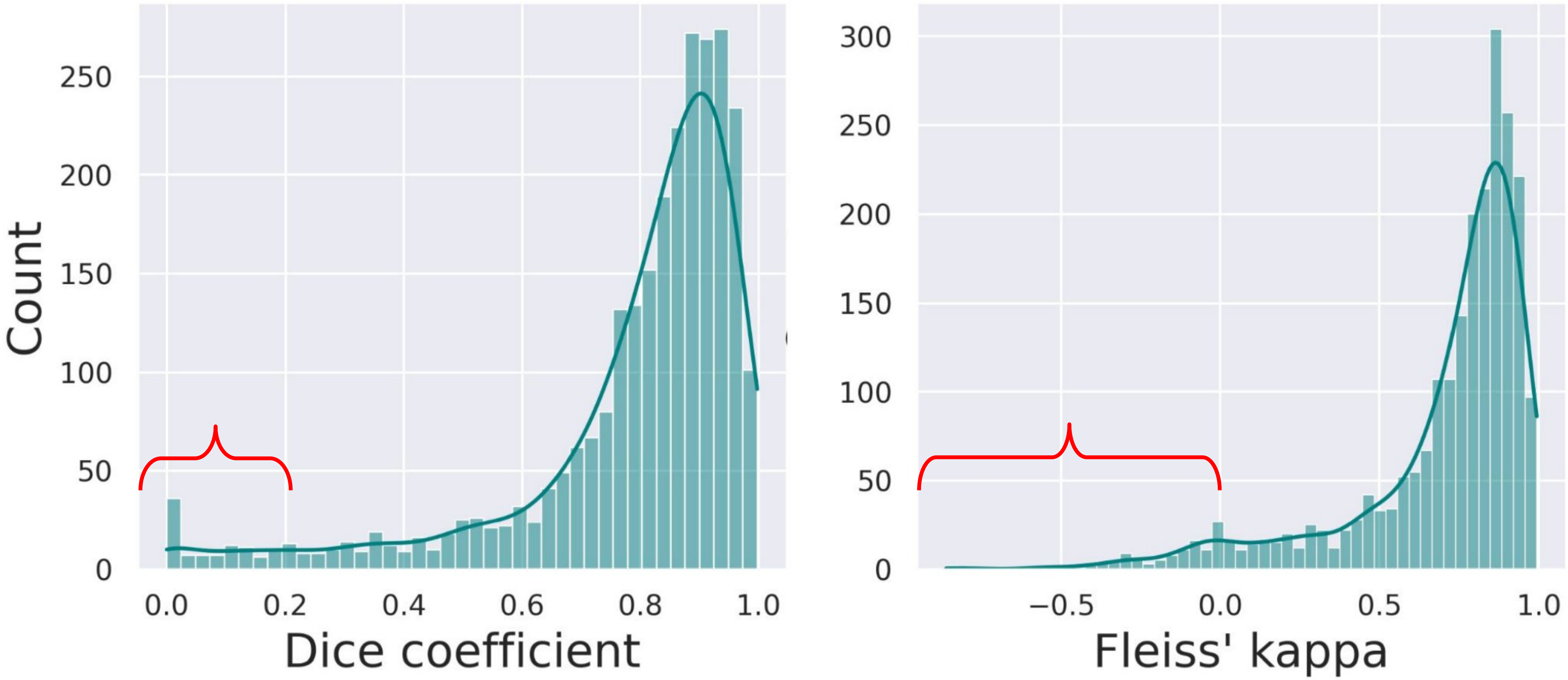




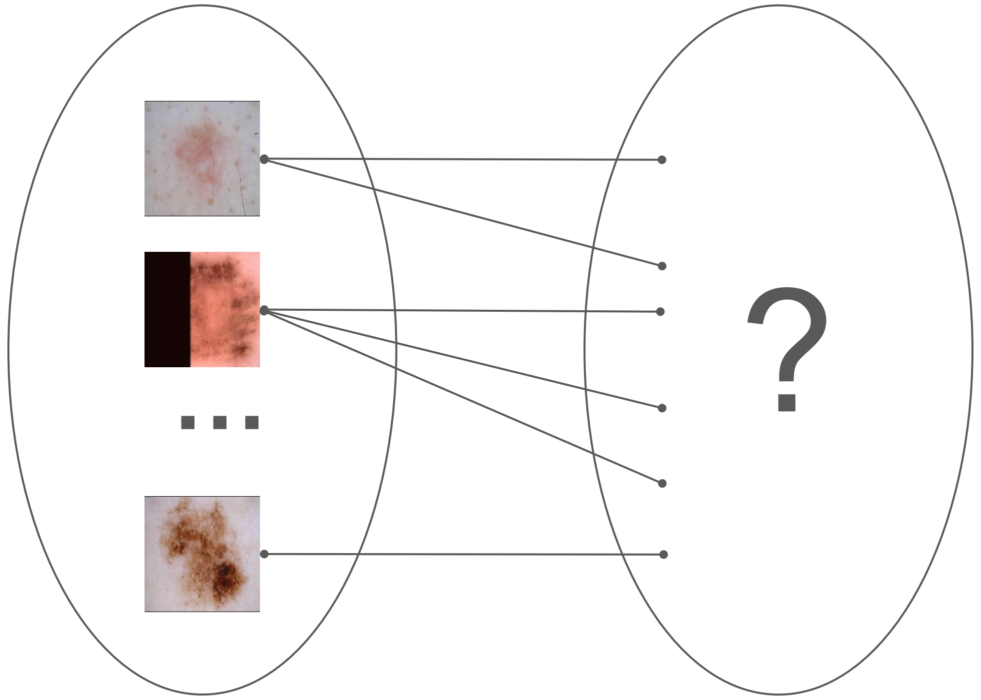
Segmentations in ISIC Archive and their Variability

2,261 images with more than 1 “ground truth” segmentation mask
⇒ 4,704 training image-mask pairs for skin lesion segmentation (**SLS**).



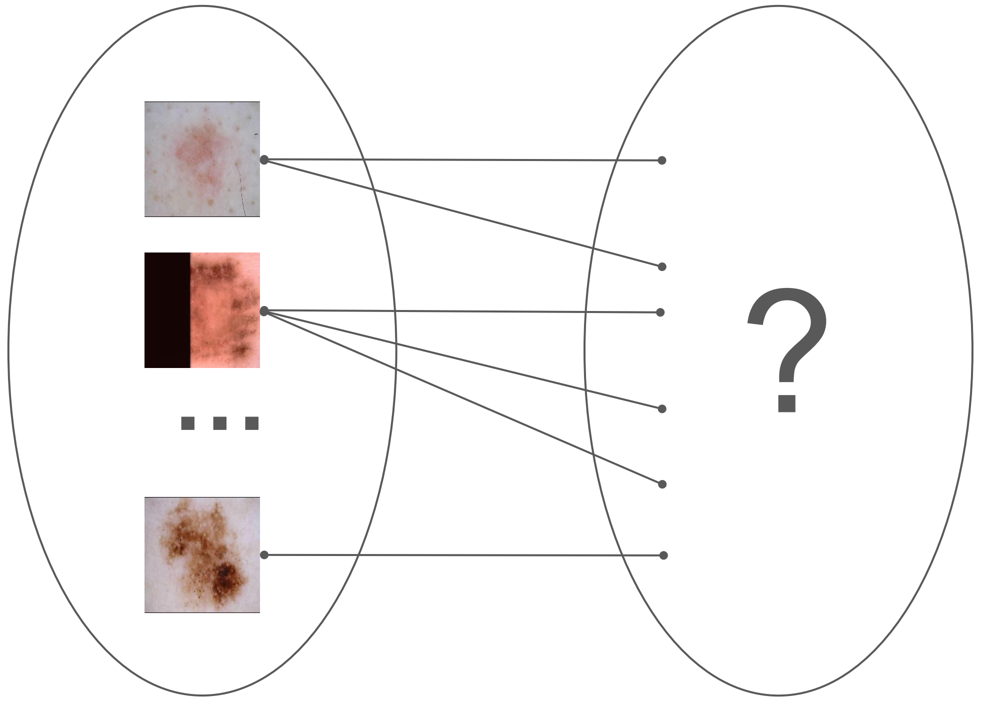


Objective

Given  , train a model that **discovers unique annotation styles** such that:

Objective

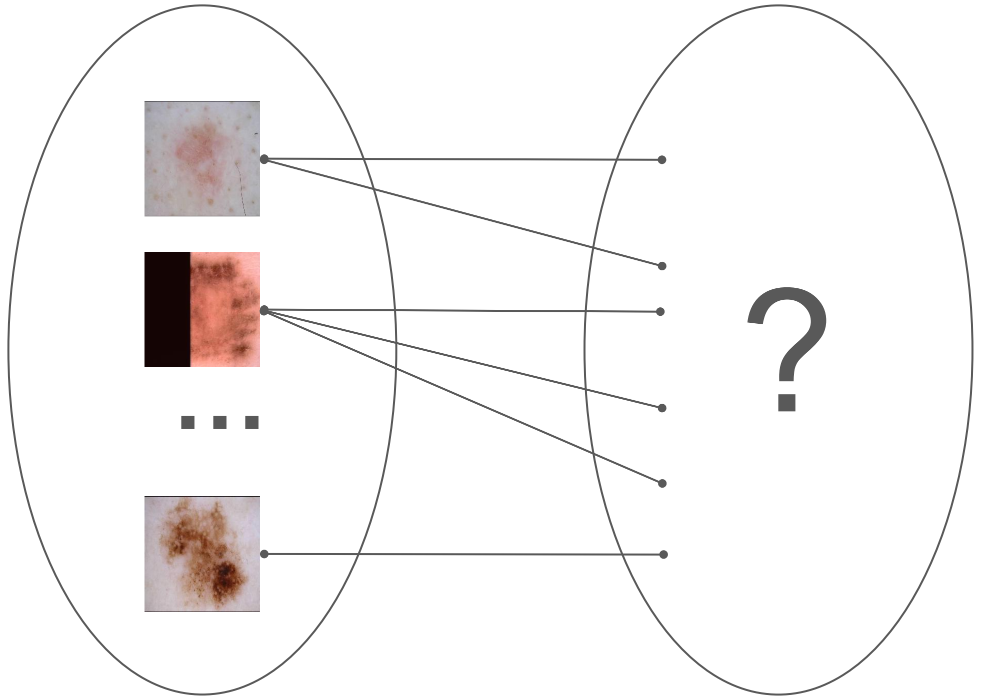
Given



, train a model that **discovers unique annotation styles** such that:

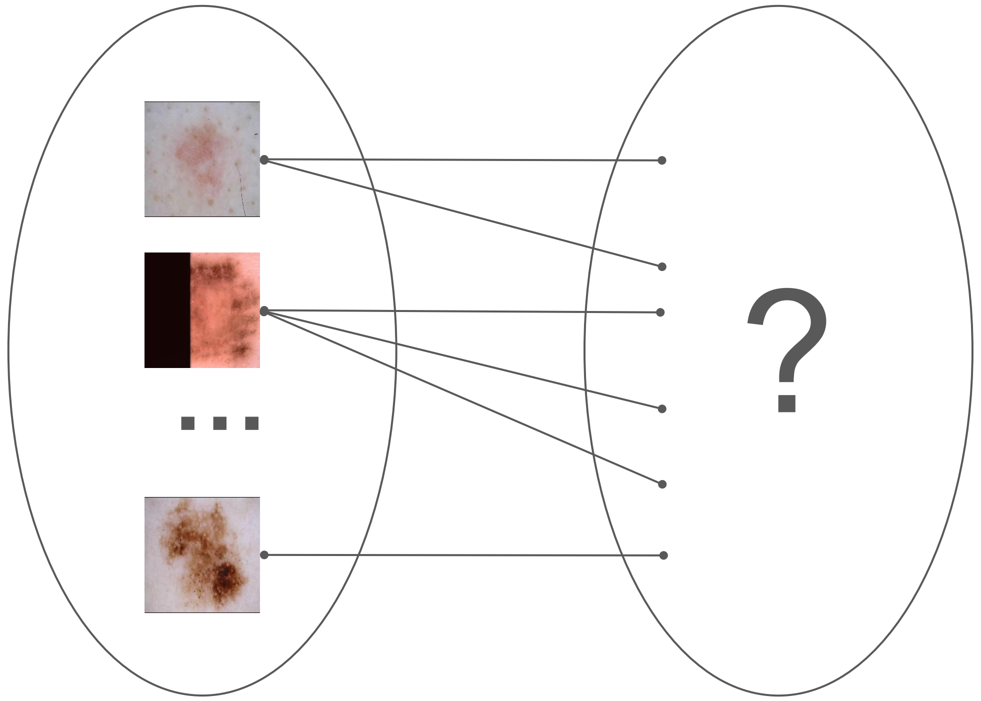
- all the predicted segmentations are **plausible**,

Objective

Given  , train a model that **discovers unique annotation styles** such that:

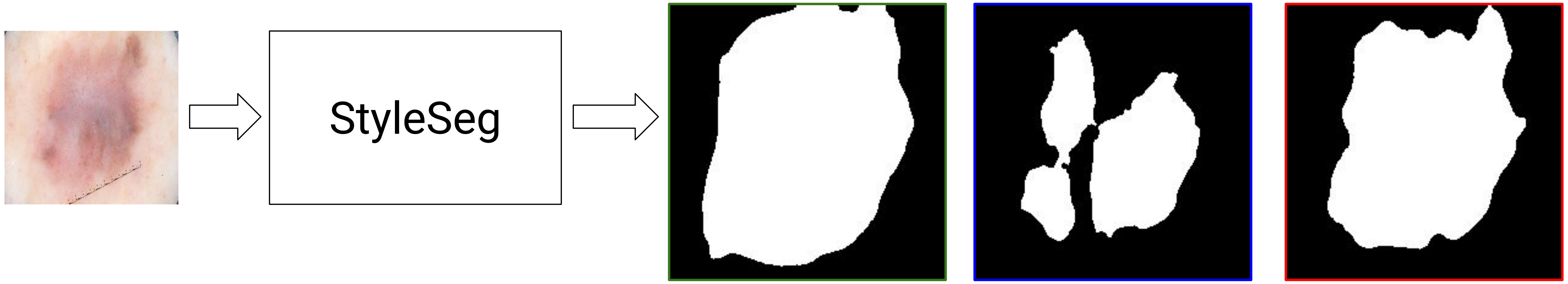
- all the predicted segmentations are **plausible**,
- the predicted segmentations are **diverse**, and

Objective

Given  , train a model that **discovers unique annotation styles** such that:

- all the predicted segmentations are **plausible**,
- the predicted segmentations are **diverse**, and
- the segmentation styles are **semantically consistent** across all images.

StyleSeg produces multiple segmentation styles



Multiple segmentation styles and their probabilities

Multiple segmentation styles and their probabilities

Style 1 Style 2 Style 3

Segmentation Model
$f_s(X_i; \Theta_s)$

M segmentation styles

Multiple segmentation styles and their probabilities

Style 1 Style 2 Style 3

Segmentation Model
$f_s(X_i; \Theta_s)$

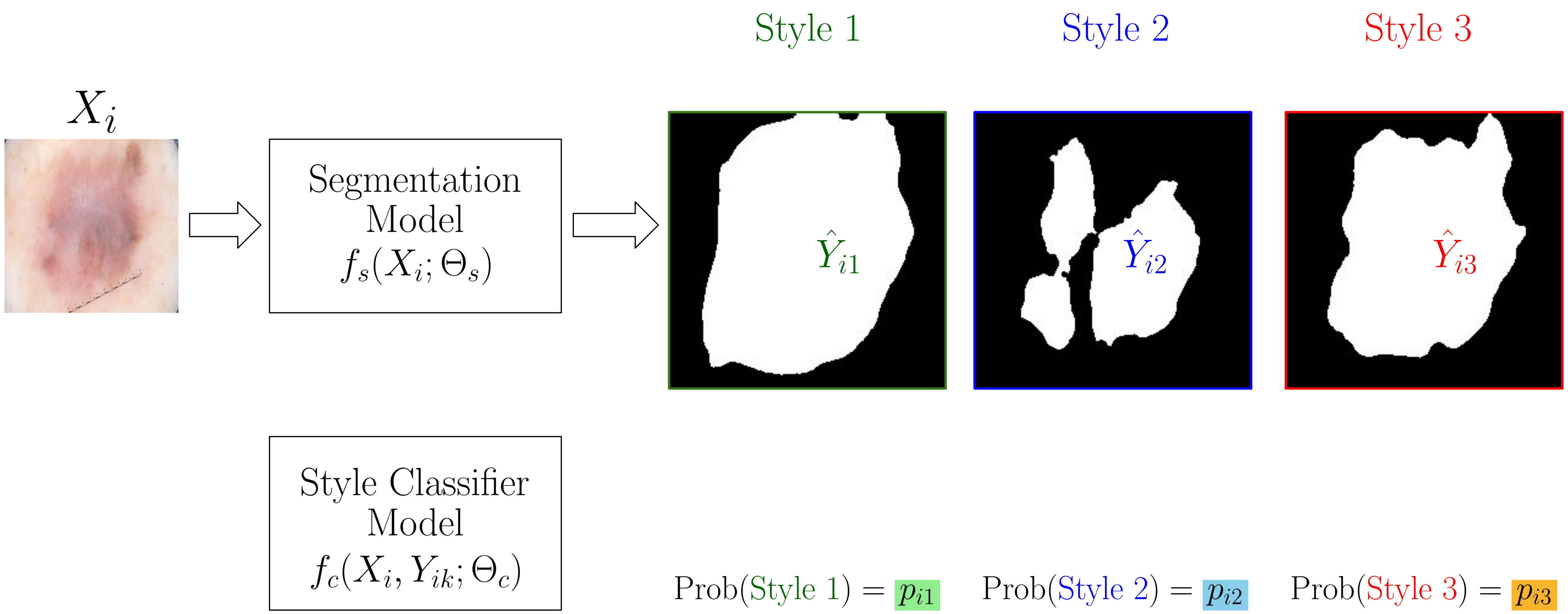
M segmentation styles

Style Classifier Model
$f_c(X_i, Y_{ik}; \Theta_c)$

$\text{Prob}(\text{Style 1}) = p_{i1}$ $\text{Prob}(\text{Style 2}) = p_{i2}$ $\text{Prob}(\text{Style 3}) = p_{i3}$

Probability of each style being the correct style.

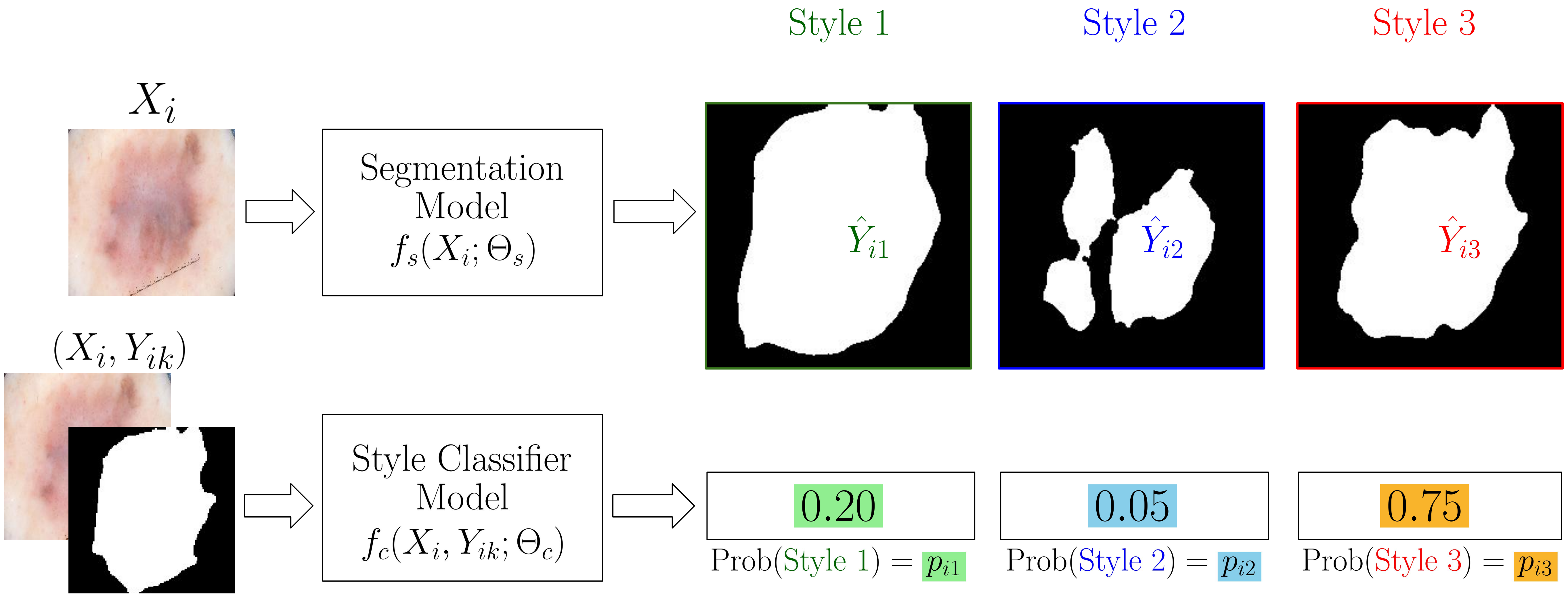
Multiple segmentation styles and their probabilities



M segmentation styles

Probability of each style being the correct style.

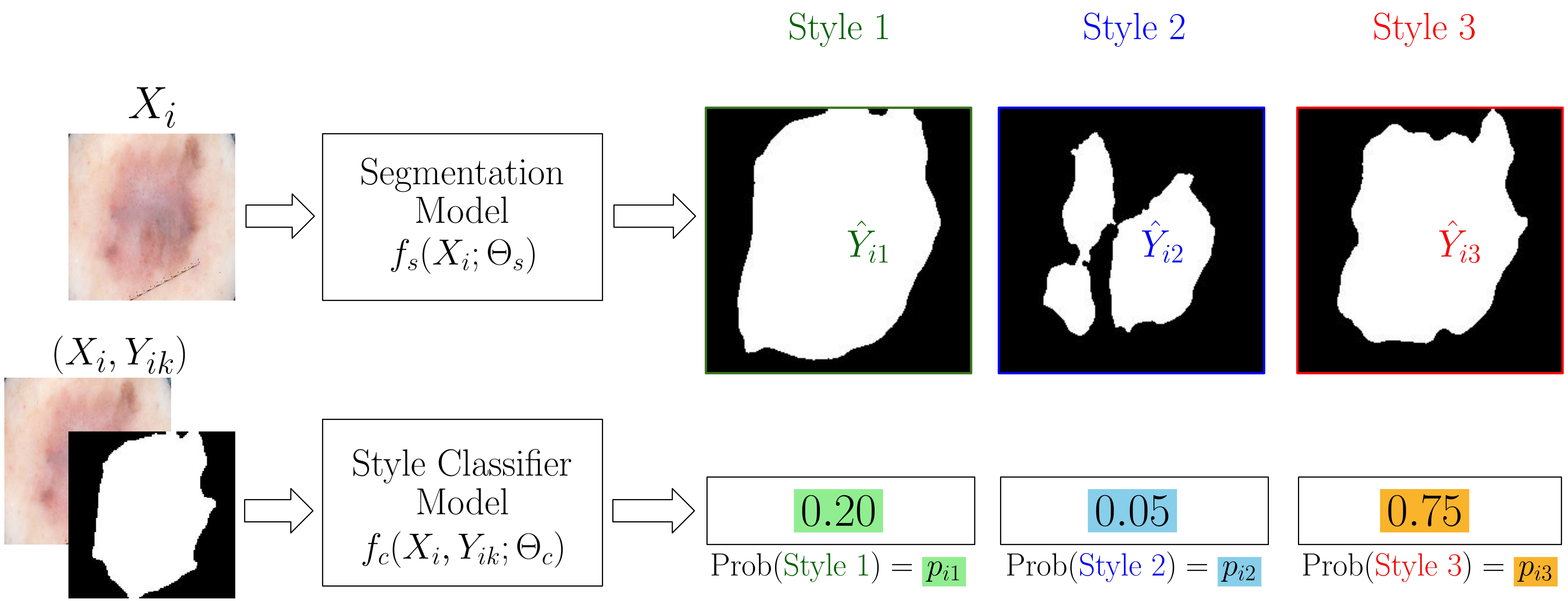
Multiple segmentation styles and their probabilities



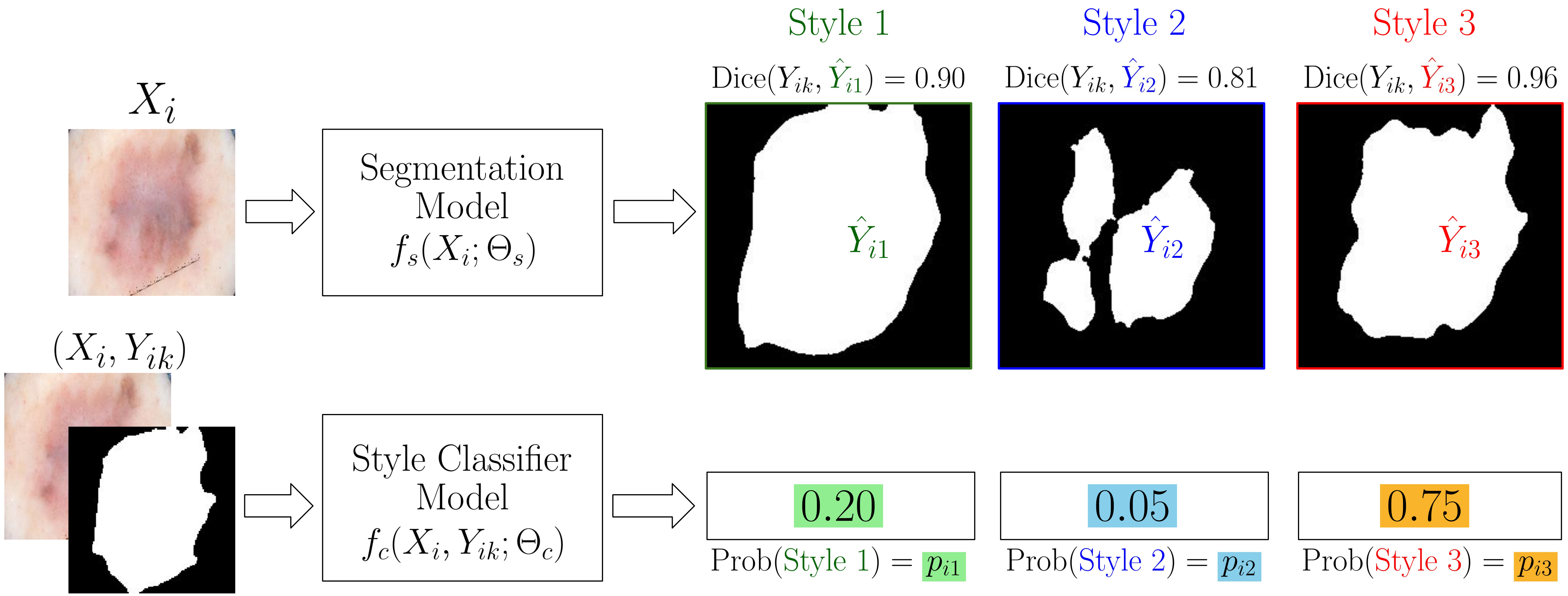
M segmentation styles

Probability of each style being the correct style.

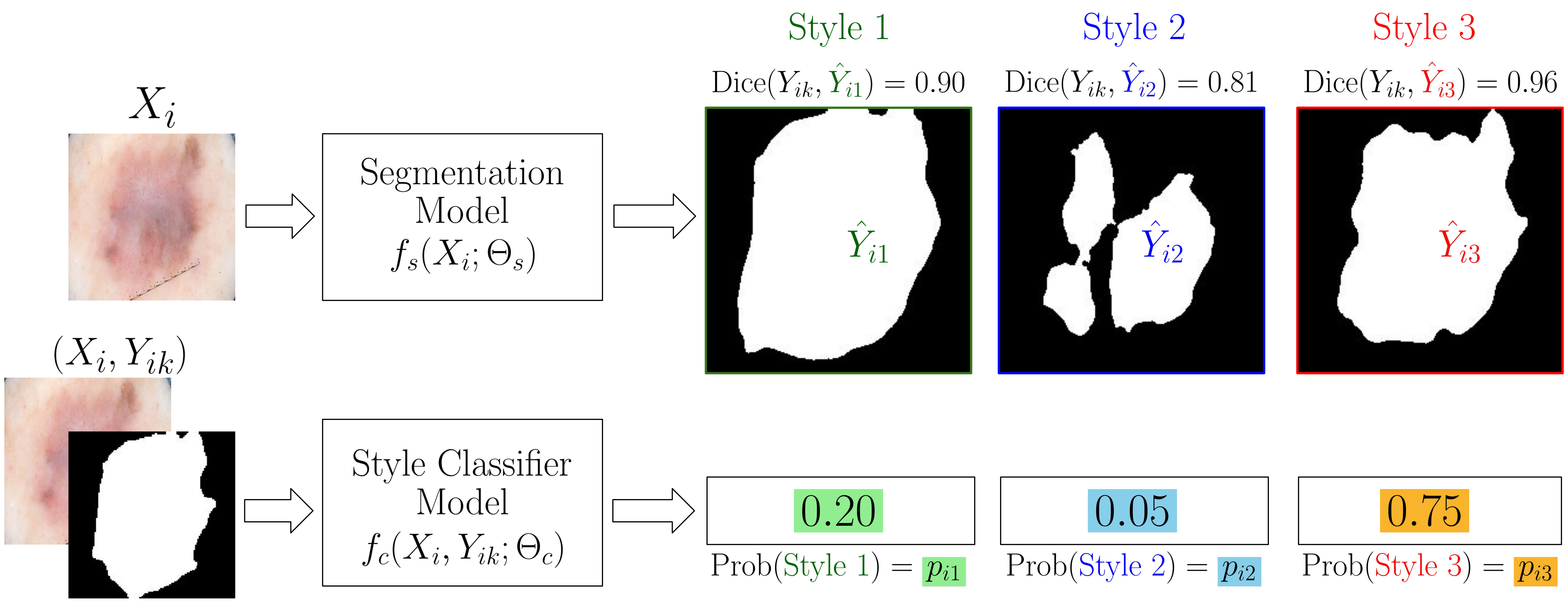
Training StyleSeg



Training StyleSeg



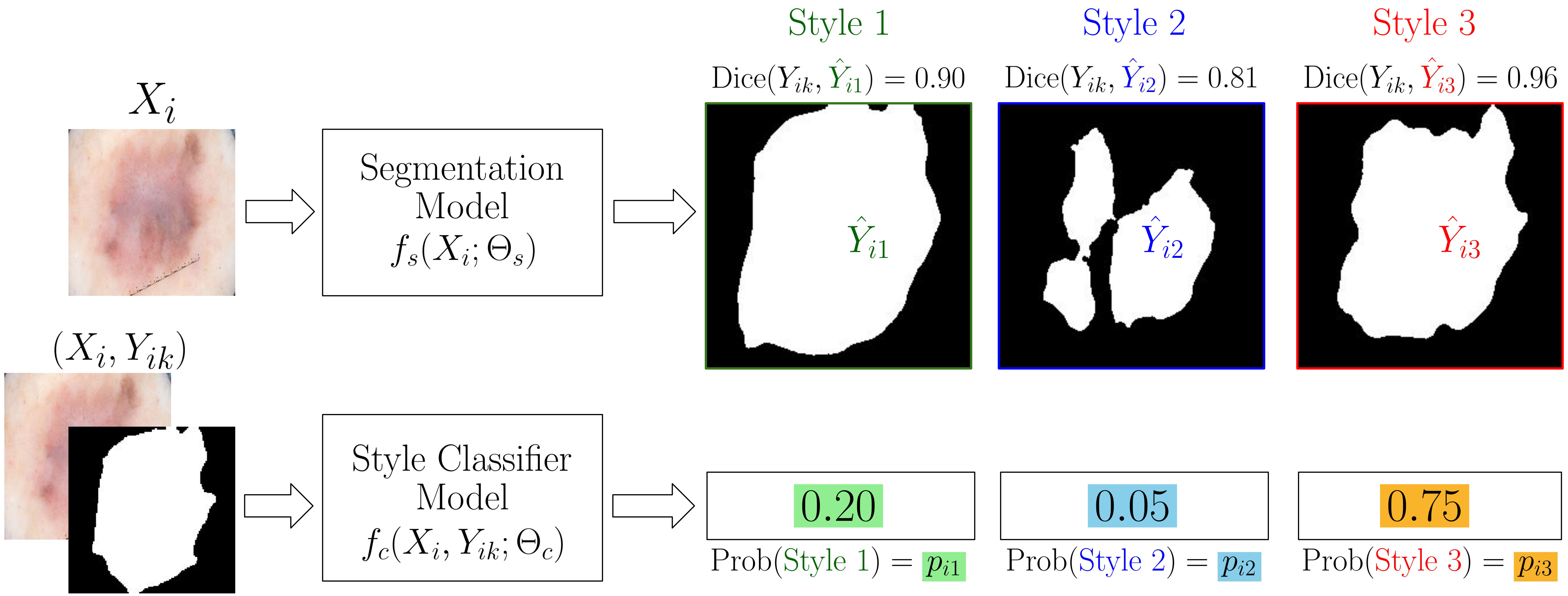
Training StyleSeg



$$m^* = \arg\max_j \text{Dice}(Y_{ik}, \hat{Y}_{ij}) = 3$$

$$\mathcal{L}_1 = L_D(Y_{ik}, \hat{Y}_{i3})$$

Training StyleSeg

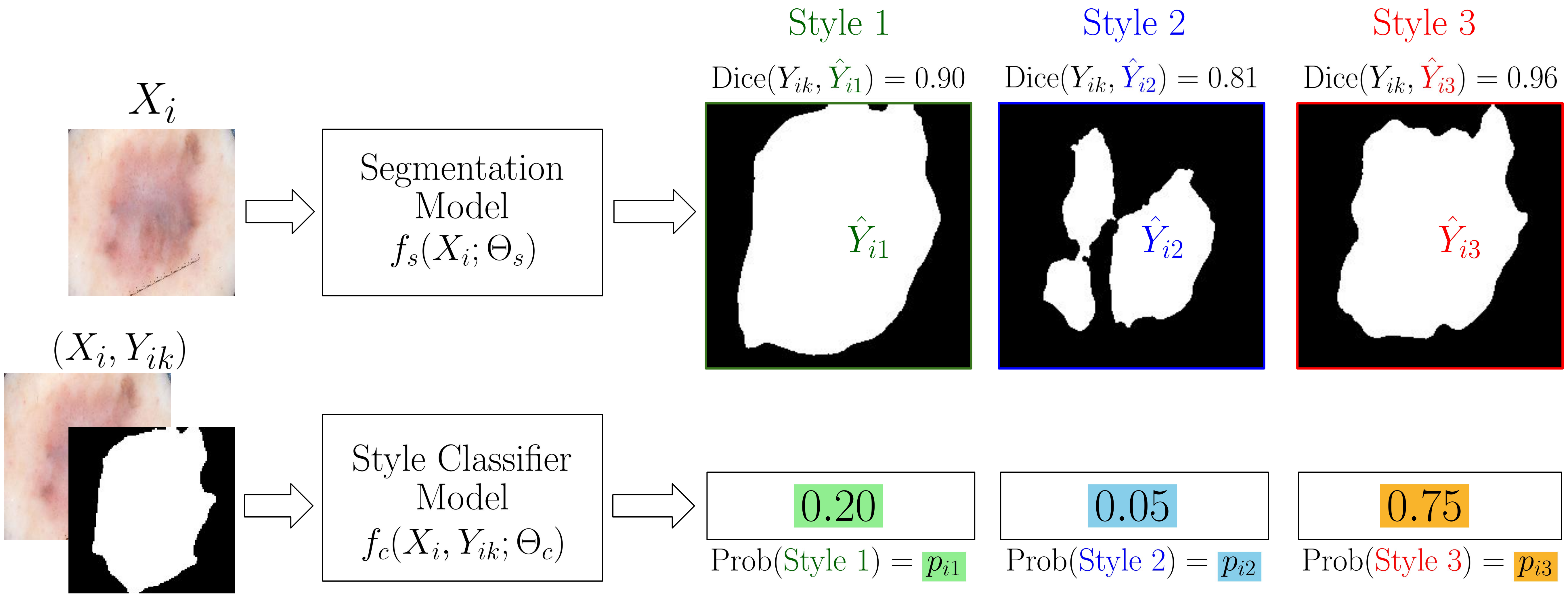


$$m^* = \arg\max_j \mathrm{Dice}(Y_{ik}, \hat{Y}_{ij}) = 3$$

$$\mathcal{L}_1 = L_D(Y_{ik}, \hat{Y}_{i3})$$

$$\mathcal{L}_2 = L_D(Y_{ik}, (0.20 * \hat{Y}_{i1}) + (0.05 * \hat{Y}_{i2}) + (0.75 * \hat{Y}_{i3}))$$

Training StyleSeg



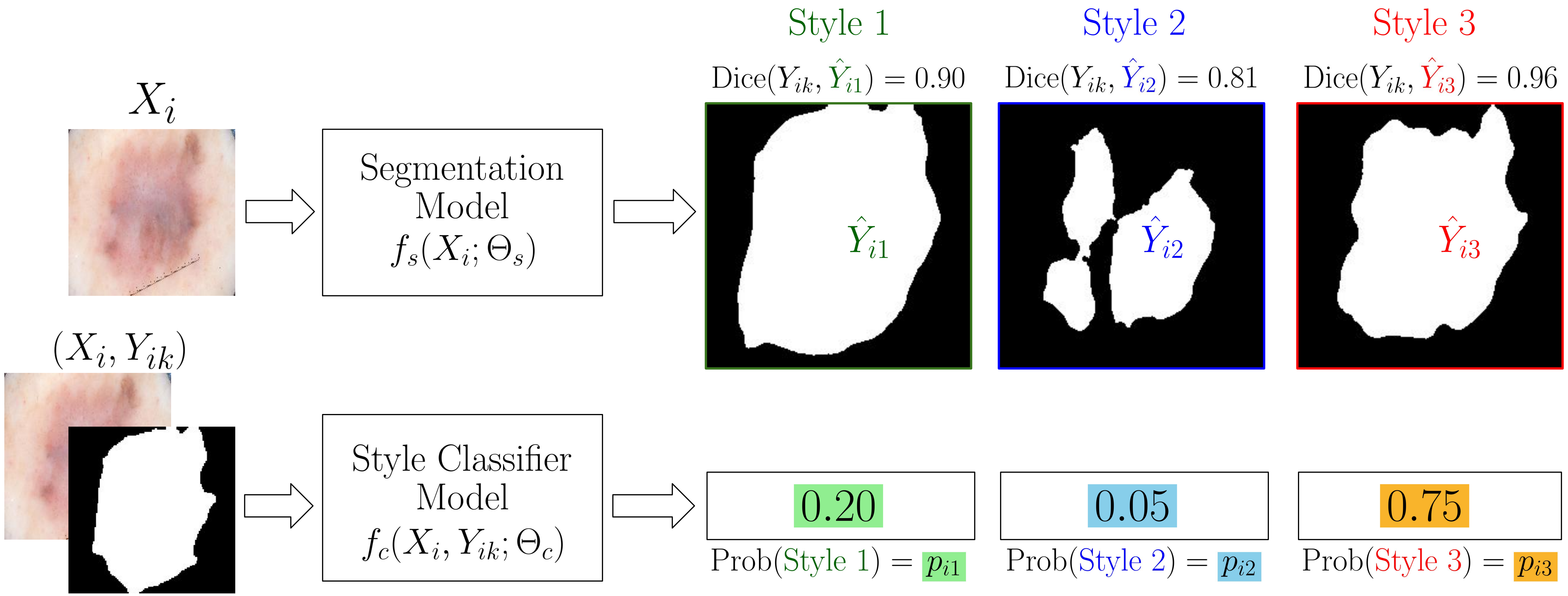
$$m^* = \arg\max_j \text{Dice}(Y_{ik}, \hat{Y}_{ij}) = 3$$

$$\mathcal{L}_1 = L_D(Y_{ik}, \hat{Y}_{i3})$$

$$\mathcal{L}_2 = L_D(Y_{ik}, (0.20 * \hat{Y}_{i1}) + (0.05 * \hat{Y}_{i2}) + (0.75 * \hat{Y}_{i3}))$$

$$\mathcal{L}_3 = L_{CE}(\ [\ 0.20,\ 0.05,\ 0.75\ ],\ [\ 0,\ 0,\ 1\ ])$$

Training StyleSeg



$$m^* = \arg\max_j \text{Dice}(Y_{ik}, \hat{Y}_{ij}) = 3$$

$$\mathcal{L}_1 = L_D(Y_{ik}, \hat{Y}_{i3})$$

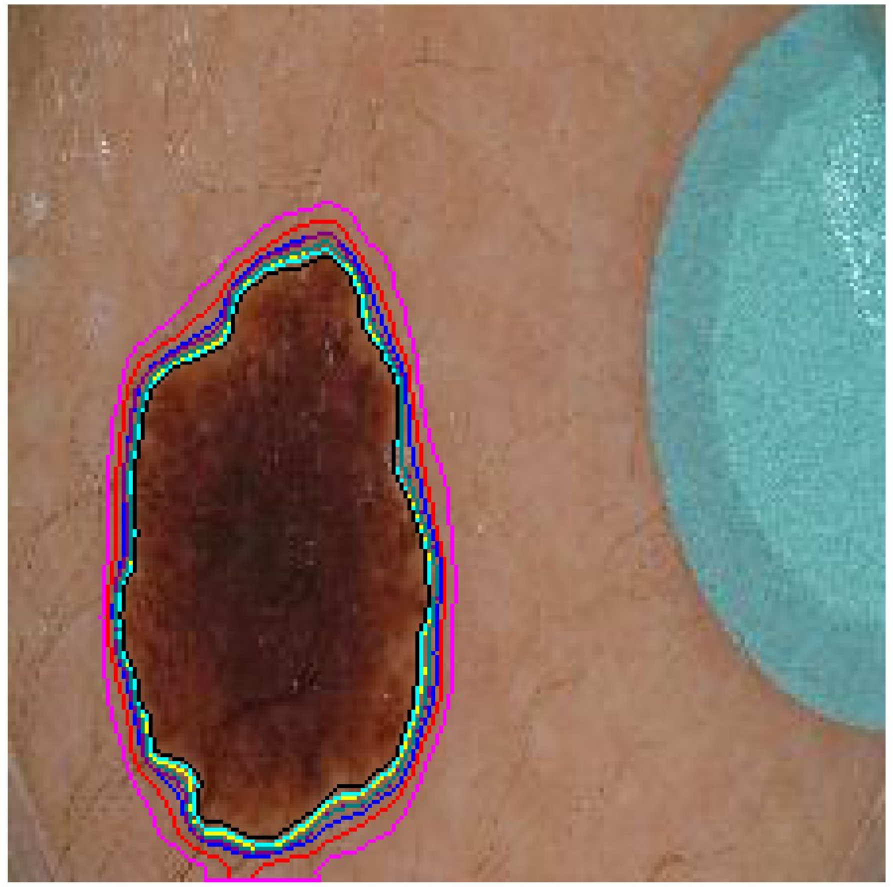
$$\mathcal{L}_2 = L_D(Y_{ik}, (0.20 * \hat{Y}_{i1}) + (0.05 * \hat{Y}_{i2}) + (0.75 * \hat{Y}_{i3}))$$

$$\mathcal{L}_3 = L_{CE}(\, [\, 0.20, 0.05, 0.75 \,], [\, 0, \; 0, \; 1 \,])$$

$$\mathcal{L}_{\text{total}} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3$$

StyleSeg Outputs Adapt to Variability in Lesion Content

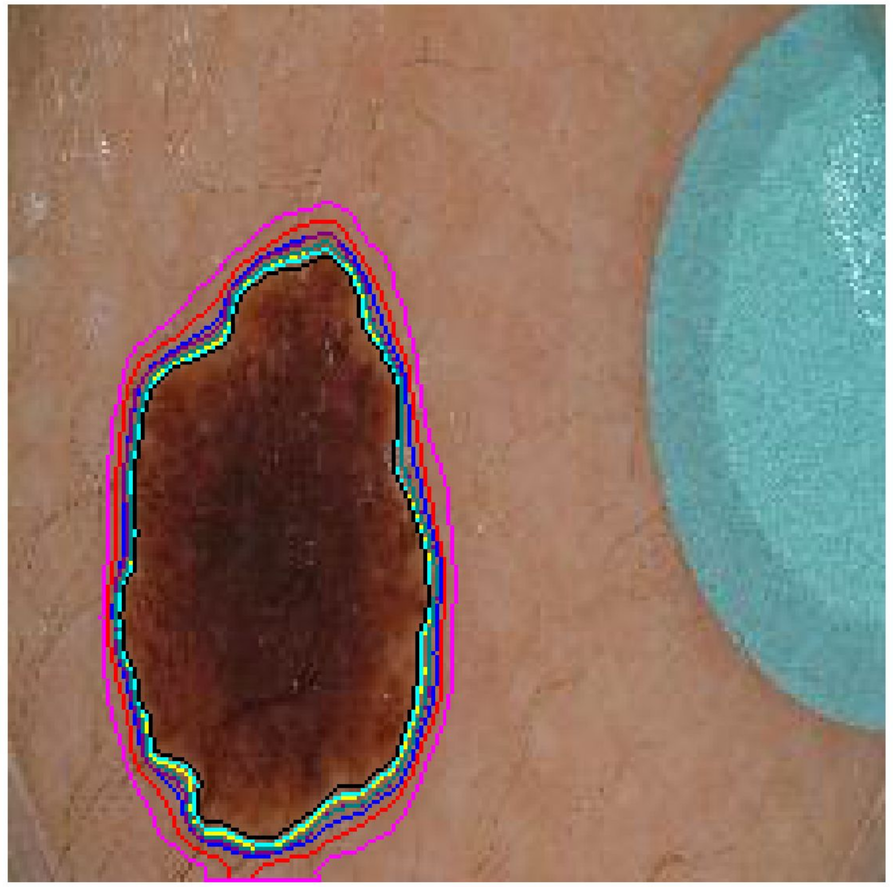
ISIC_0003599



High-contrast lesion has **high agreement** across styles

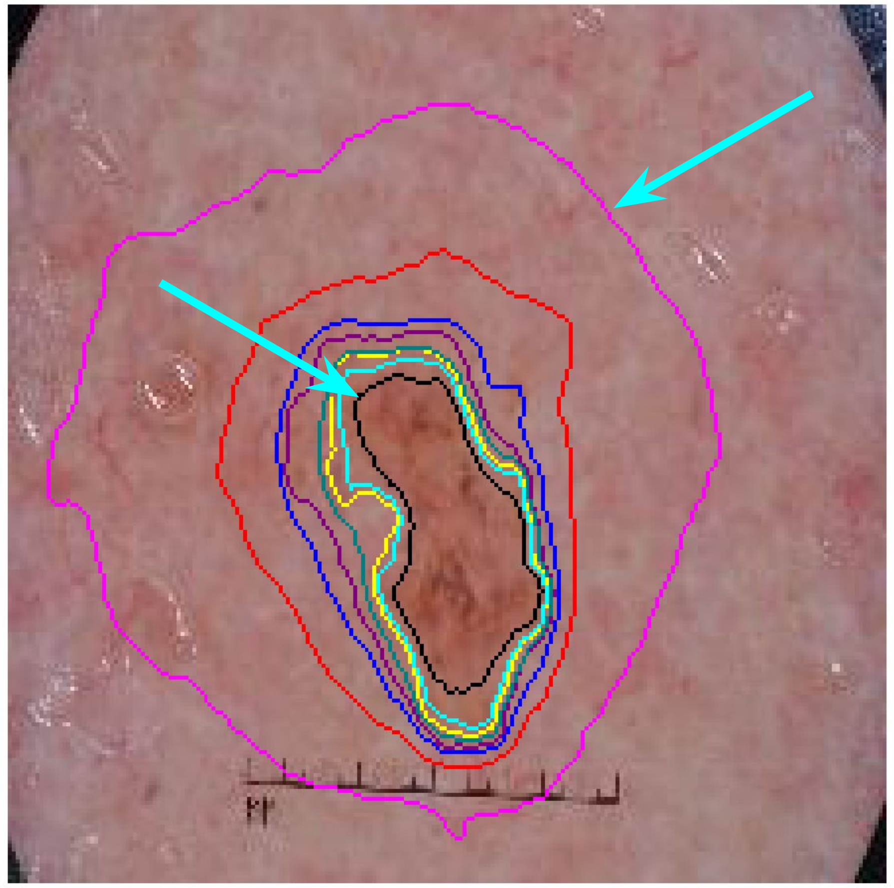
StyleSeg Outputs Adapt to Variability in Lesion Content

ISIC_0003599



High-contrast lesion has **high agreement** across styles

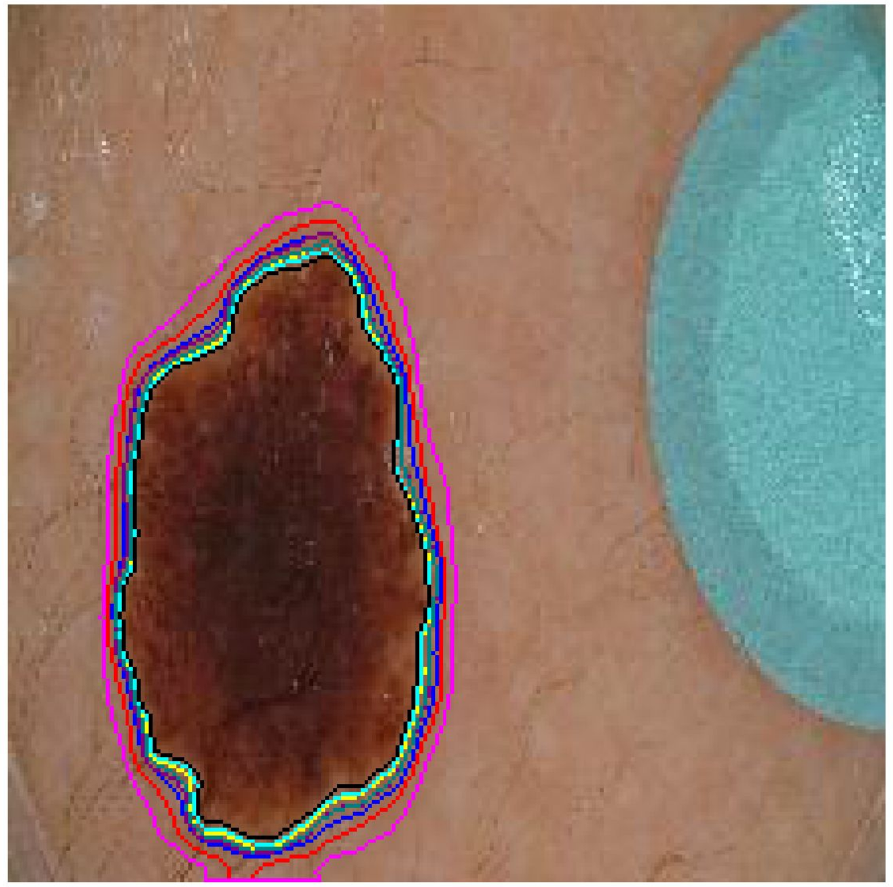
ISIC_0014337



Instances of **under-** and **over-** segmentation

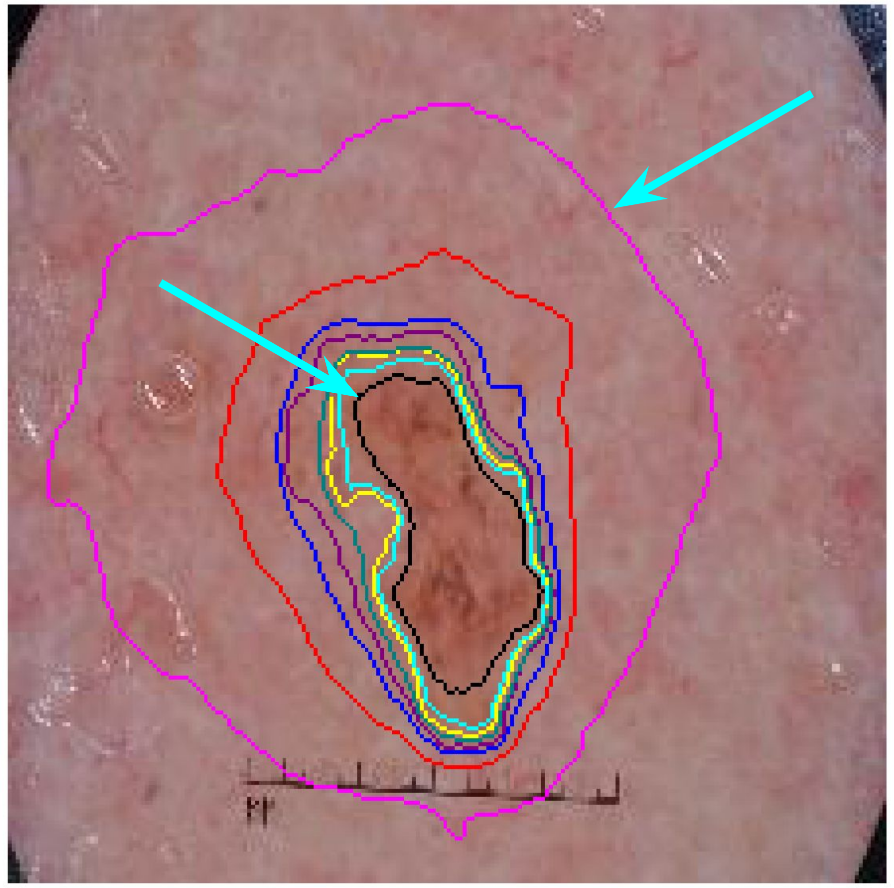
StyleSeg Outputs Adapt to Variability in Lesion Content

ISIC_0003599



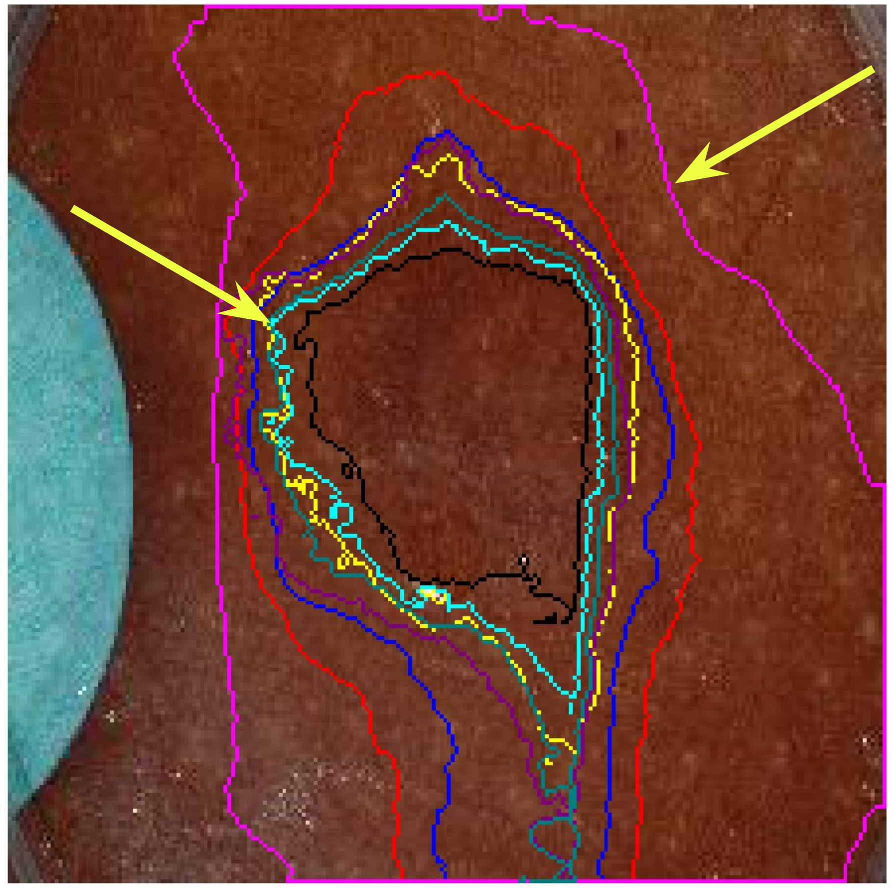
High-contrast lesion has **high agreement** across styles

ISIC_0014337



Instances of **under-** and **over-** segmentation

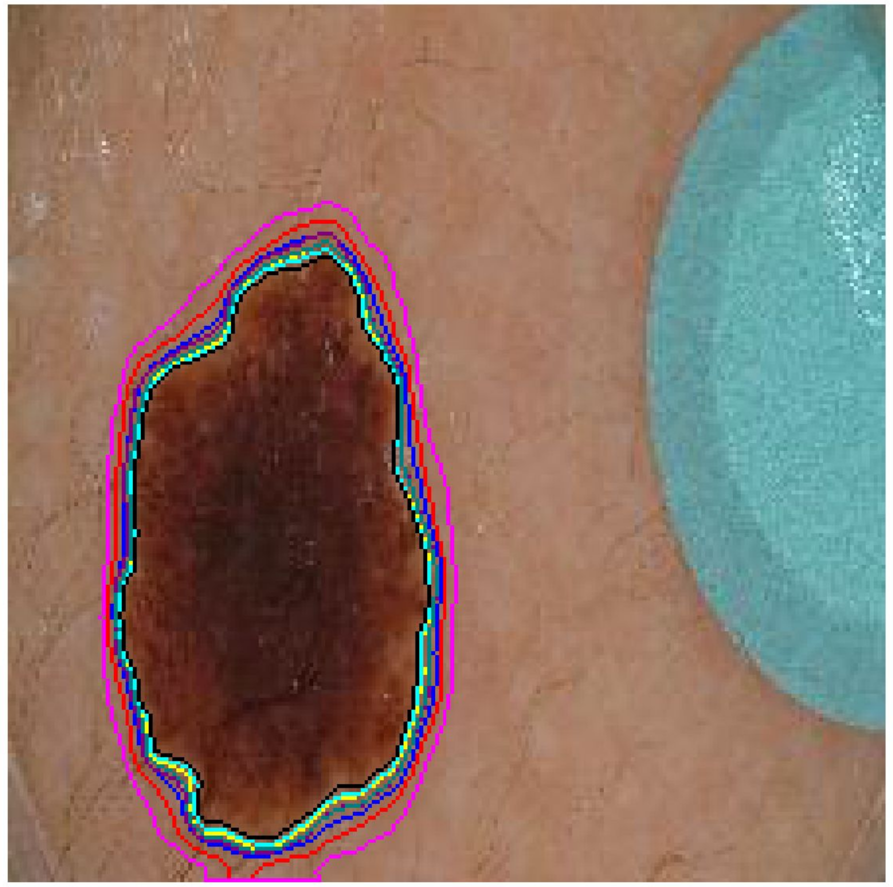
ISIC_0003726



Different **boundary jaggedness** across segmentations

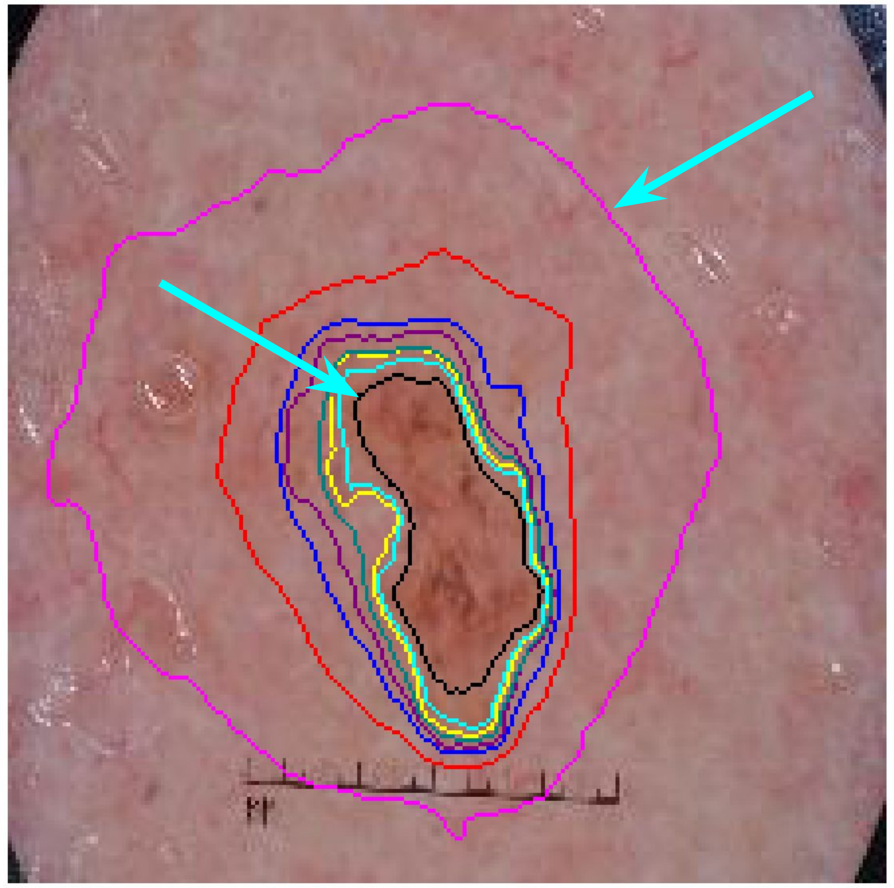
StyleSeg Outputs Adapt to Variability in Lesion Content

ISIC_0003599



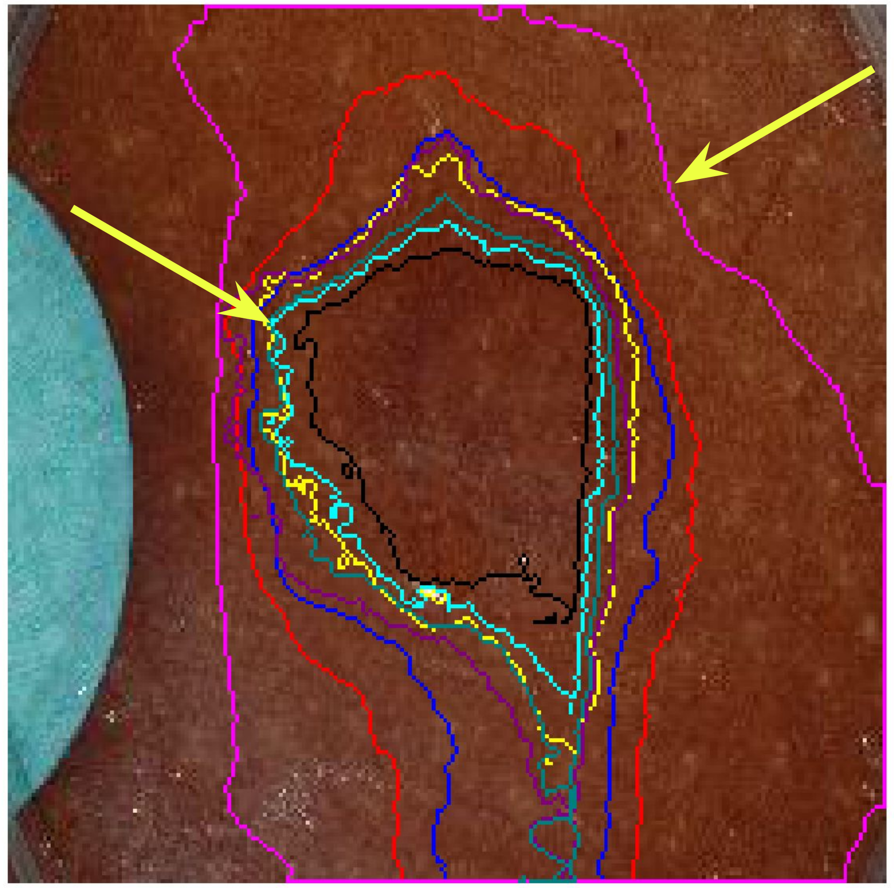
High-contrast lesion has **high agreement** across styles

ISIC_0014337



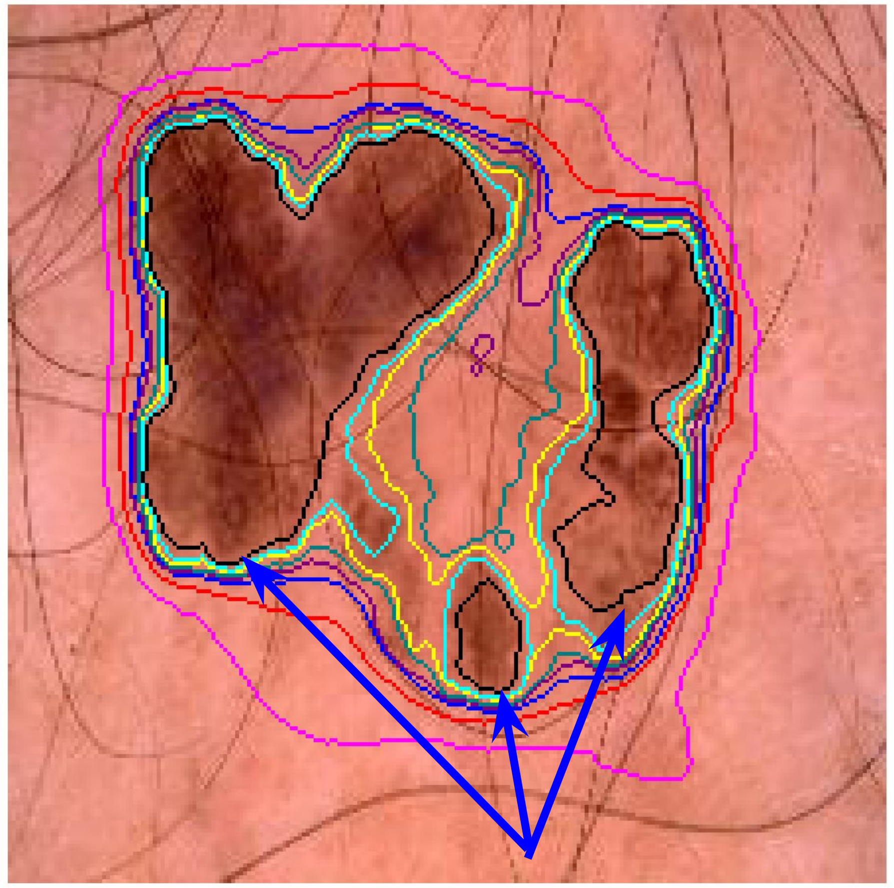
Instances of **under-** and **over-** segmentation

ISIC_0003726



Different **boundary jaggedness** across segmentations

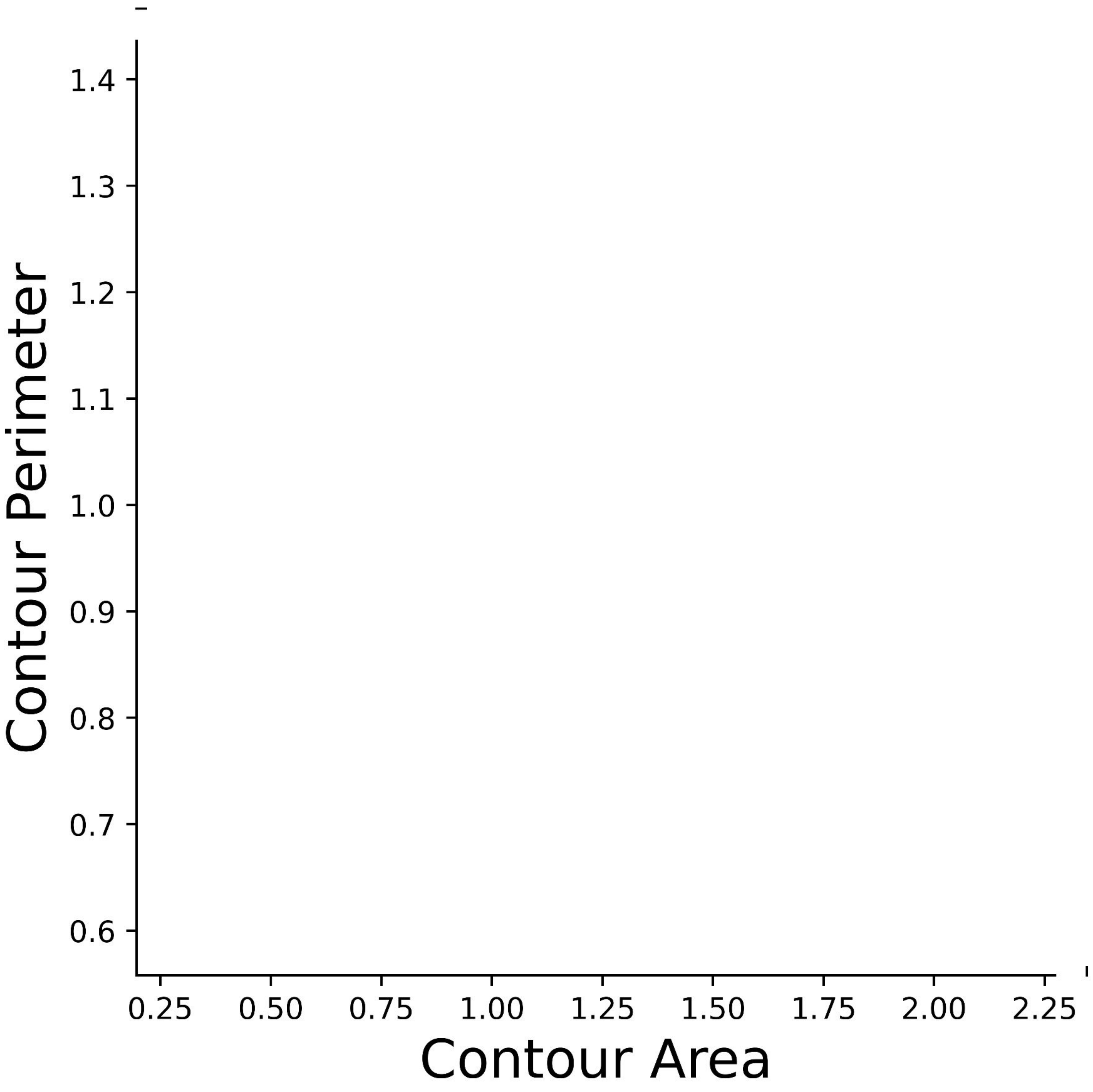
ISIC_0014831



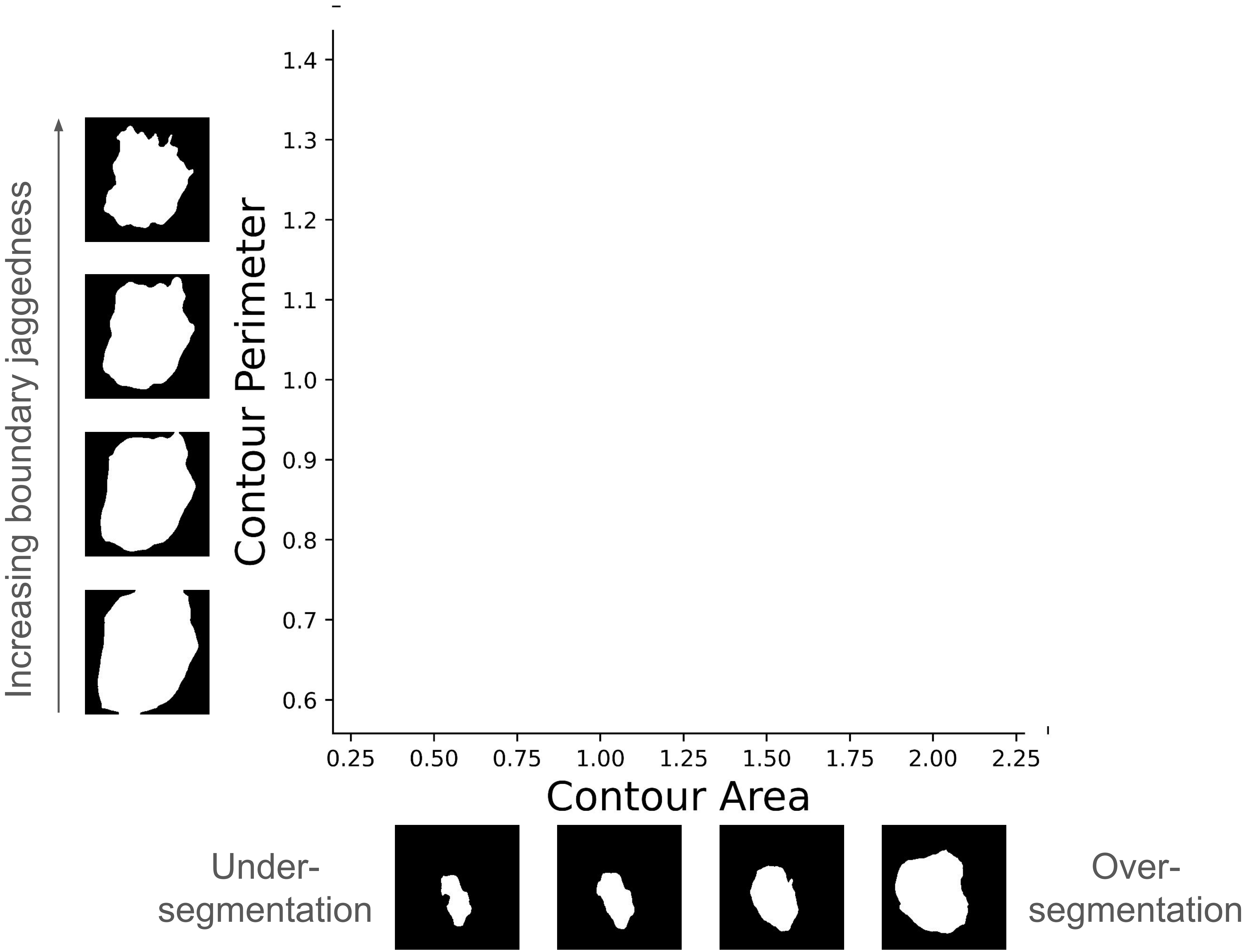
Ambiguous boundary causes segmentation masks to **split**

Semantic Consistency of Styles

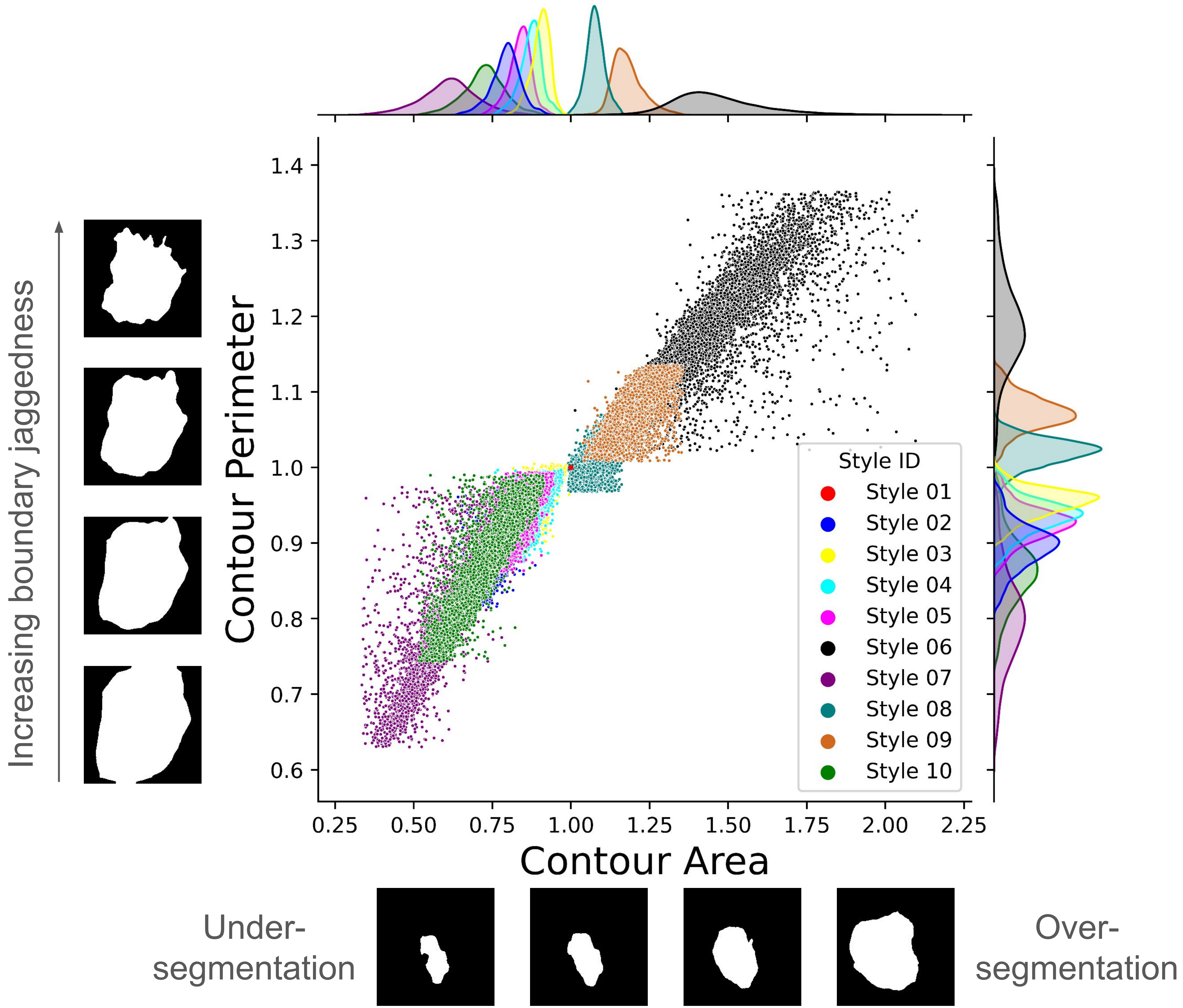
Semantic Consistency of Styles



Semantic Consistency of Styles



Semantic Consistency of Styles



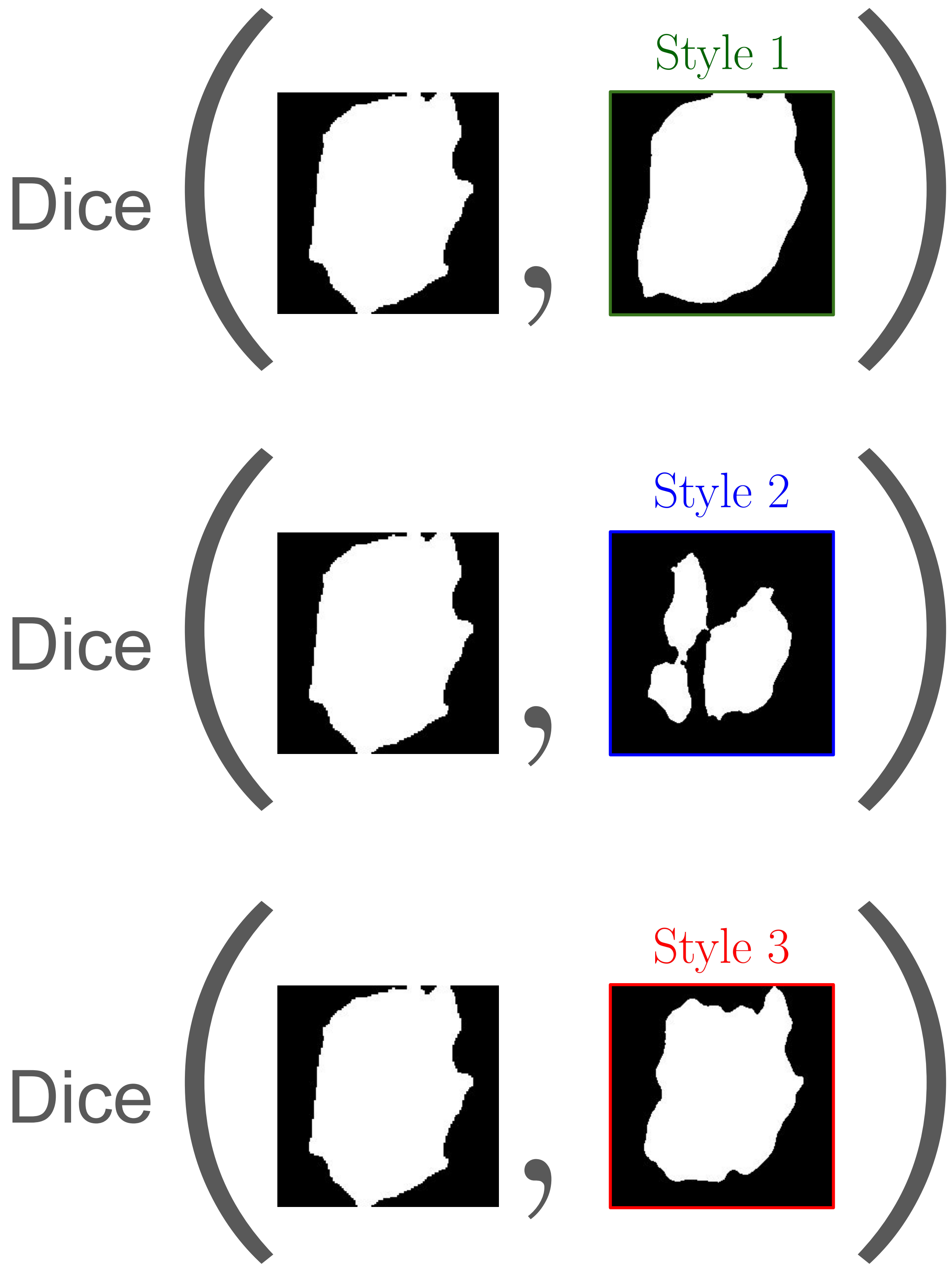
Competing Methods

$\mathbb{SS}\mathrm{eg}$ methods	
NaiveTraining	SLS model without any annotator-specific knowledge.
RandAnnotID[2]	4 SLS models, one optimized for each annotator randomly assigned to a mask.
LessIsMore[3]	SLS model trained on a subset of the masks whose average pairwise Cohen's kappa ≥ 0.5.
D-LEMA[2]	Ensemble of Bayesian SLS models.

Competing Methods

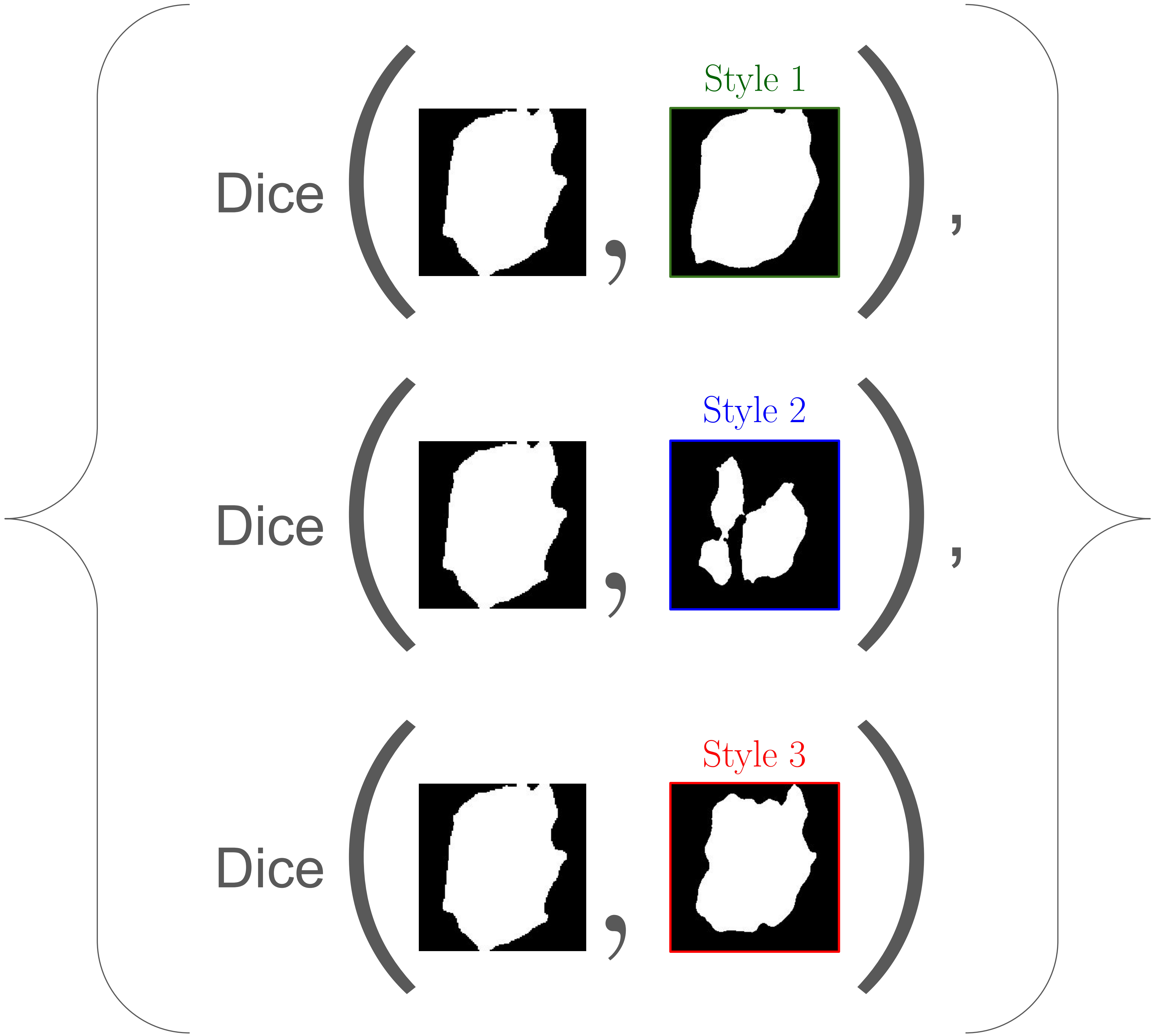
SSeg methods	
NaiveTraining	SLS model without any annotator-specific knowledge.
RandAnnotID[2]	4 SLS models, one optimized for each annotator randomly assigned to a mask.
LessIsMore[3]	SLS model trained on a subset of the masks whose average pairwise Cohen's kappa ≥ 0.5.
D-LEMA[2]	Ensemble of Bayesian SLS models.
MSeg **methods**	
MHP[4]	Multi-hypothesis prediction model, repurposed for SLS.

Metrics



Metrics

- min.
- max.



Quantitative Results

Method	ISIC Archive-Test ($n = 10,000$)		DermoFit ($n = 1,300$)	
	Min. Dice	Max. Dice	Min. Dice	Max. Dice
NaiveTraining	0.800		0.842	
RandAnnotID	–		0.826	
LessIsMore	0.815		0.854	
D-LEMA	–		0.853	
2-MHP	0.727	0.864	0.707	0.882
2-StyleSeg	0.760	0.869	0.759	0.888
3-MHP	0.652	0.876	0.562	0.888
3-StyleSeg	0.713	0.881	0.720	0.897
4-MHP	0.623	0.886	0.636	0.904
4-StyleSeg	0.693	0.889	0.681	0.907
6-MHP	0.121	0.886	0.428	0.900
6-StyleSeg	0.648	0.889	0.651	0.911
8-MHP	0.099	0.896	0.309	0.908
8-StyleSeg	0.595	0.899	0.632	0.910
10-MHP	0.281	0.894	0.181	0.906
10-StyleSeg	0.603	0.899	0.579	0.918

Results on 4 datasets:

- **ISIC Archive-Test** (n = 10000)
- **DermoFit** (n = 1300)
- PH^2 (n = 200)
- SCD (n = 206)

Learning Multiple Styles Is Always Better

Method	ISIC Archive-Test ($n = 10,000$)		DermoFit ($n = 1,300$)	
	Min. Dice	Max. Dice	Min. Dice	Max. Dice
NaiveTraining	0.800		0.842	
RandAnnotID	–		0.826	
LessIsMore	0.815		0.854	
D-LEMA	–		0.853	
2-MHP	0.727	0.864	0.707	0.882
2-StyleSeg	0.760	0.869	0.759	0.888
3-MHP	0.652	0.876	0.562	0.888
3-StyleSeg	0.713	0.881	0.720	0.897
4-MHP	0.623	0.886	0.636	0.904
4-StyleSeg	0.693	0.889	0.681	0.907
6-MHP	0.121	0.886	0.428	0.900
6-StyleSeg	0.648	0.889	0.651	0.911
8-MHP	0.099	0.896	0.309	0.908
8-StyleSeg	0.595	0.899	0.632	0.910
10-MHP	0.281	0.894	0.181	0.906
10-StyleSeg	0.603	0.899	0.579	0.918

Learning to predict more than 1 style (MSeg methods), <u>even learning to predict 2 styles</u>, consistently outperforms SSeg methods.

Diversity Increases As More Styles are Learned

Method	**ISIC Archive-Test** ($n = 10,000$)		**DermoFit** ($n = 1,300$)	
	Min. Dice	Max. Dice	Min. Dice	Max. Dice
NaiveTraining	0.800		0.842	
RandAnnotID	–		0.826	
LessIsMore	0.815		0.854	
D-LEMA	–		0.853	
2-MHP	0.727	0.864	0.707	0.882
2-StyleSeg	0.760	$\boxed{0.869}$	0.759	$\boxed{0.888}$
3-MHP	0.652	0.876	0.562	0.888
3-StyleSeg	0.713	$\boxed{0.881}$	0.720	$\boxed{0.897}$
4-MHP	0.623	0.886	0.636	0.904
4-StyleSeg	0.693	$\boxed{0.889}$	0.681	$\boxed{0.907}$
6-MHP	0.121	0.886	0.428	0.900
6-StyleSeg	0.648	$\boxed{0.889}$	0.651	$\boxed{0.911}$
8-MHP	0.099	0.896	0.309	0.908
8-StyleSeg	0.595	$\boxed{0.899}$	0.632	$\boxed{0.910}$
10-MHP	0.281	0.894	0.181	0.906
10-StyleSeg	0.603	$\boxed{0.899}$	0.579	$\boxed{0.918}$

As *M* increases, a larger number of diverse segmentations are generated, and the max. Dice keeps improving.

StyleSeg Outperforms MHP

Method	ISIC Archive-Test ($n = 10,000$)		DermoFit ($n = 1,300$)	
	Min. Dice	Max. Dice	Min. Dice	Max. Dice
NaiveTraining	0.800		0.842	
RandAnnotID	–		0.826	
LessIsMore	0.815		0.854	
D-LEMA	–		0.853	
2-MHP	0.727	0.864	0.707	0.882
2-StyleSeg	0.760	0.869	0.759	0.888
3-MHP	0.652	0.876	0.562	0.888
3-StyleSeg	0.713	0.881	0.720	0.897
4-MHP	0.623	0.886	0.636	0.904
4-StyleSeg	0.693	0.889	0.681	0.907
6-MHP	0.121	0.886	0.428	0.900
6-StyleSeg	0.648	0.889	0.651	0.911
8-MHP	0.099	0.896	0.309	0.908
8-StyleSeg	0.595	0.899	0.632	0.910
10-MHP	0.281	0.894	0.181	0.906
10-StyleSeg	0.603	0.899	0.579	0.918

StyleSeg consistently outperforms MHP for all values of *M* and for all datasets.

StyleSeg Outputs Are More Plausible

Method	ISIC Archive-Test ($n = 10,000$)		DermoFit ($n = 1,300$)	
	Min. Dice	Max. Dice	Min. Dice	Max. Dice
NaiveTraining	0.800		0.842	
RandAnnotID	–		0.826	
LessIsMore	0.815		0.854	
D-LEMA	–		0.853	
2-MHP	0.727	0.864	0.707	0.882
2-StyleSeg	0.760	0.869	0.759	0.888
3-MHP	0.652	0.876	0.562	0.888
3-StyleSeg	0.713	0.881	0.720	0.897
4-MHP	0.623	0.886	0.636	0.904
4-StyleSeg	0.693	0.889	0.681	0.907
6-MHP	0.121	0.886	0.428	0.900
6-StyleSeg	0.648	0.889	0.651	0.911
8-MHP	0.099	0.896	0.309	0.908
8-StyleSeg	0.595	0.899	0.632	0.910
10-MHP	0.281	0.894	0.181	0.906
10-StyleSeg	0.603	0.899	0.579	0.918

StyleSeg consistently outperforms MHP for all values of *M* and for all datasets.

Moreover, as M increases, all StyleSeg outputs remain reasonably plausible, whereas MHP outputs exhibit diversity at the cost of plausibility.

Performance Improves Even on Single Annot. Datasets

Method	ISIC Archive-Test ($n = 10,000$)		DermoFit ($n = 1,300$)	
	Min. Dice	Max. Dice	Min. Dice	Max. Dice
NaiveTraining	0.800		0.842	
RandAnnotID	–		0.826	
LessIsMore	0.815		0.854	
D-LEMA	–		0.853	
2-MHP	0.727	0.864	0.707	0.882
2-StyleSeg	0.760	0.869	0.759	0.888
3-MHP	0.652	0.876	0.562	0.888
3-StyleSeg	0.713	0.881	0.720	0.897
4-MHP	0.623	0.886	0.636	0.904
4-StyleSeg	0.693	0.889	0.681	0.907
6-MHP	0.121	0.886	0.428	0.900
6-StyleSeg	0.648	0.889	0.651	0.911
8-MHP	0.099	0.896	0.309	0.908
8-StyleSeg	0.595	0.899	0.632	0.910
10-MHP	0.281	0.894	0.181	0.906
10-StyleSeg	0.603	0.899	0.579	0.918

Even for datasets without documented variability in segmentations, learning to predict multiple styles is helpful.

Performance Improves Even on Single Annot. Datasets

Method	ISIC Archive-Test ($n = 10,000$)		DermoFit ($n = 1,300$)	
	Min. Dice	Max. Dice	Min. Dice	Max. Dice
NaiveTraining	0.800		0.842	
RandAnnotID	–		0.826	
LessIsMore	0.815		0.854	
D-LEMA	–		0.853	
2-MHP	0.727	0.864	0.707	0.882
2-StyleSeg	0.760	0.869	0.759	0.888
3-MHP	0.652	0.876	0.562	0.888
3-StyleSeg	0.713	0.881	0.720	0.897
4-MHP	0.623	0.886	0.636	0.904
4-StyleSeg	0.693	0.889	0.681	0.907
6-MHP	0.121	0.886	0.428	0.900
6-StyleSeg	0.648	0.889	0.651	0.911
8-MHP	0.099	0.896	0.309	0.908
8-StyleSeg	0.595	0.899	0.632	0.910
10-MHP	0.281	0.894	0.181	0.906
10-StyleSeg	0.603	0.899	0.579	0.918

Even for datasets without documented variability in segmentations, learning to predict multiple styles is helpful.

A New Multi-Annotator SLS Dataset: ISIC-MultiAnnot

The **largest** multi-annotator SLS dataset curated from the ISIC Archive.

A New Multi-Annotator SLS Dataset: ISIC-MultiAnnot

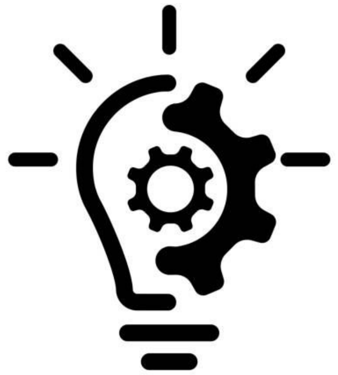
The **largest** multi-annotator SLS dataset curated from the ISIC Archive.

12,951 images

×

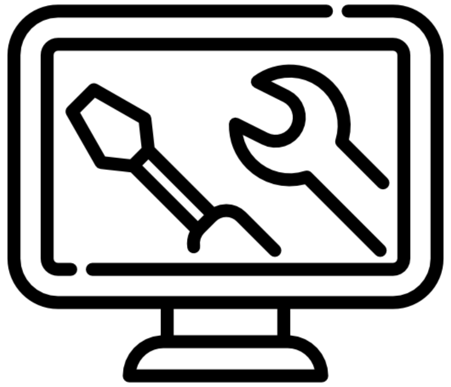
10 anonymized annotators
“A00” – “A09”

×



2 skill levels
“expert”, “novice”

×



3 tool choices
“T1” – “T3”

A New Multi-Annotator SLS Dataset: <u>ISIC-MultiAnnot</u>

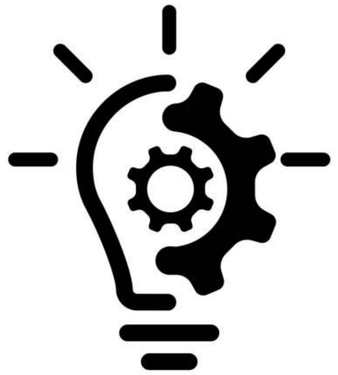
The **largest** multi-annotator SLS dataset curated from the ISIC Archive.

12,951 images

×

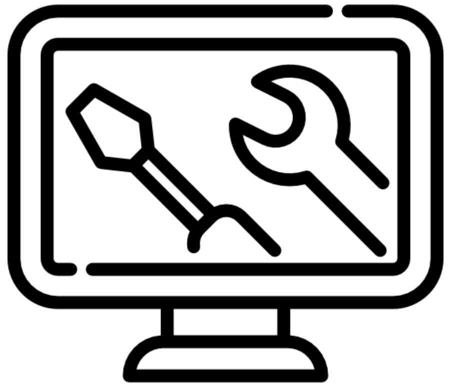
10 anonymized annotators
“A00” – “A09”

×



2 skill levels
“expert”, “novice”

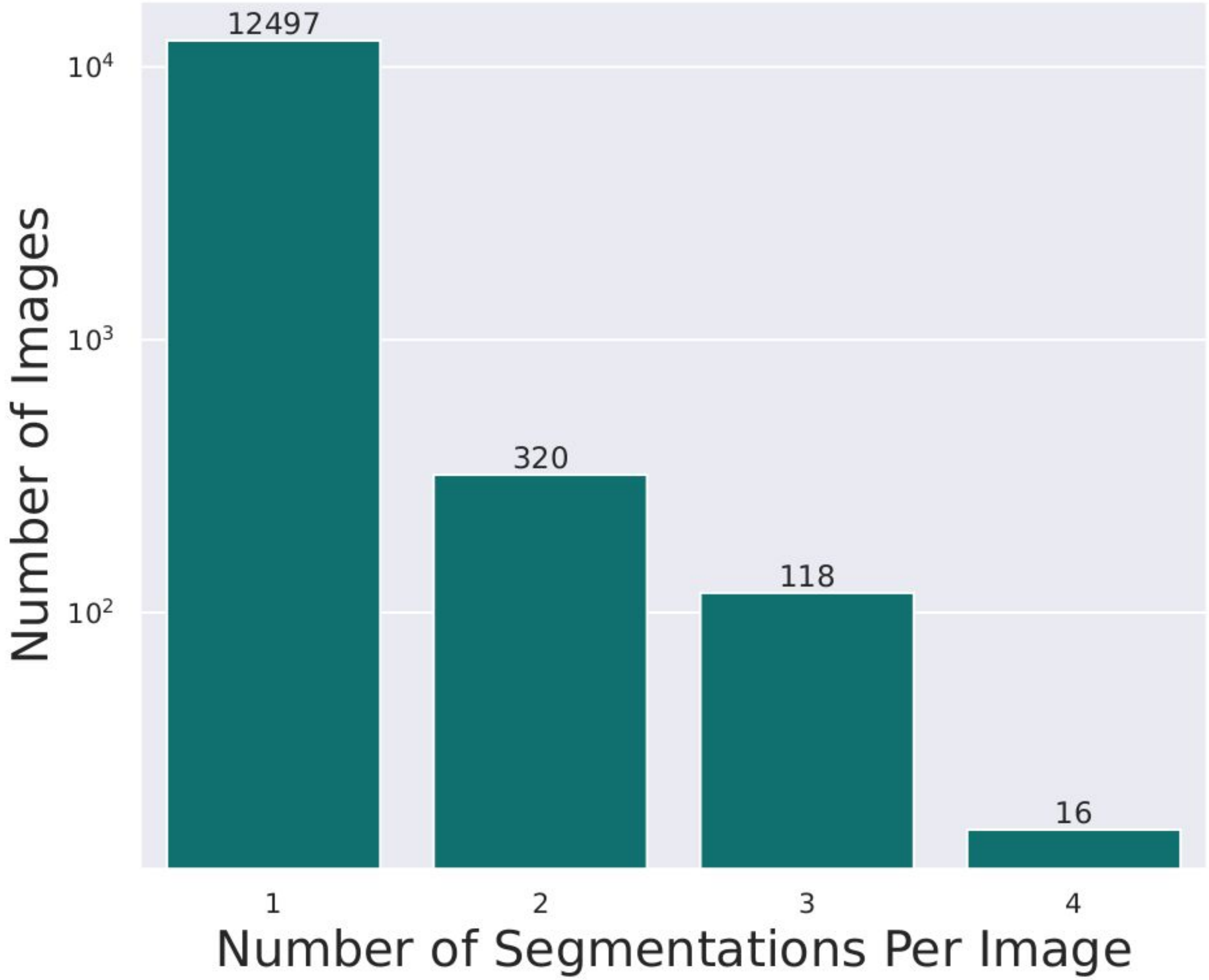
×

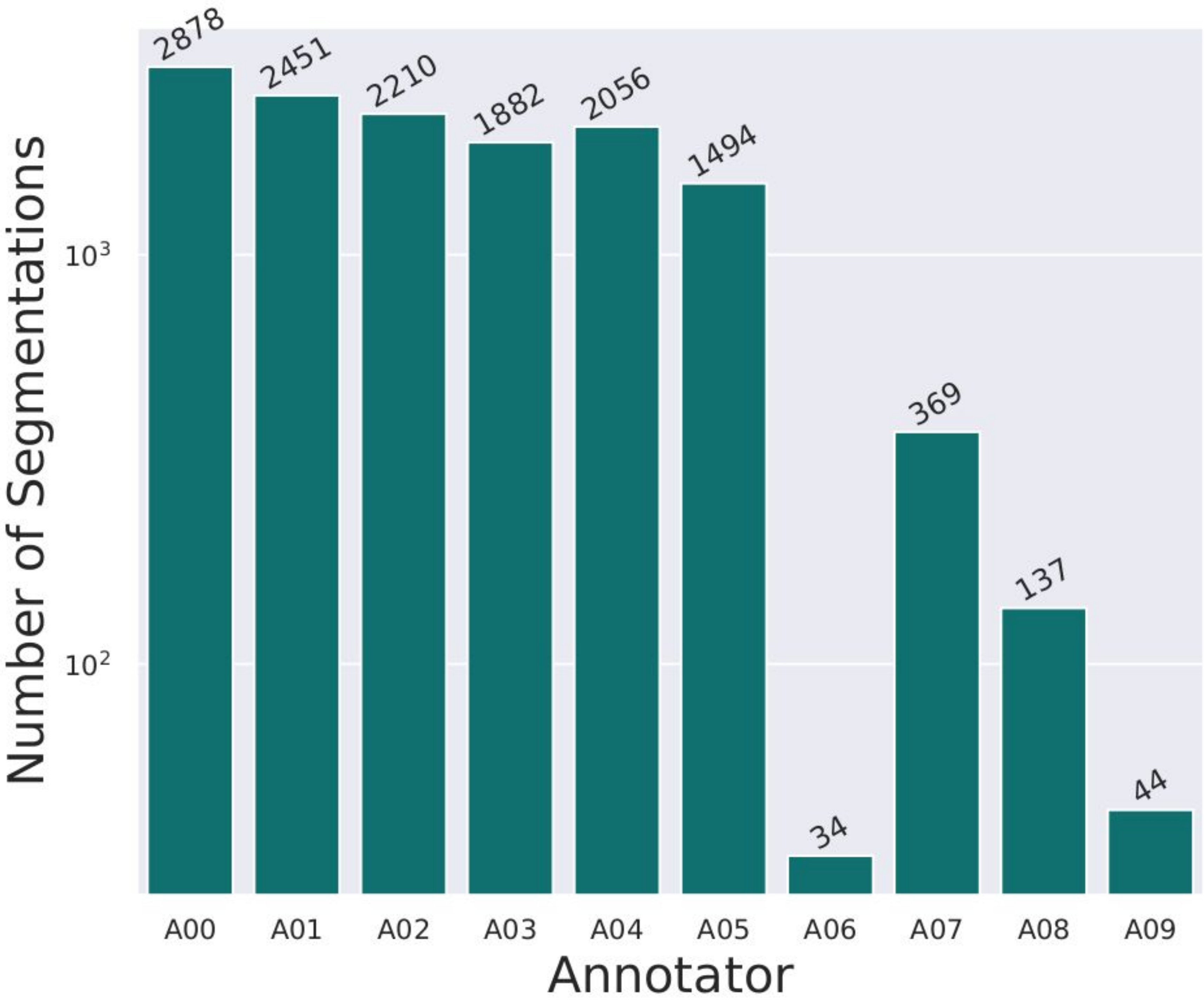


3 tool choices
“T1” – “T3”

13,555 image-mask pairs
27 unique annotator preferences

A New Multi-Annotator SLS Dataset: <u>ISIC-MultiAnnot</u>

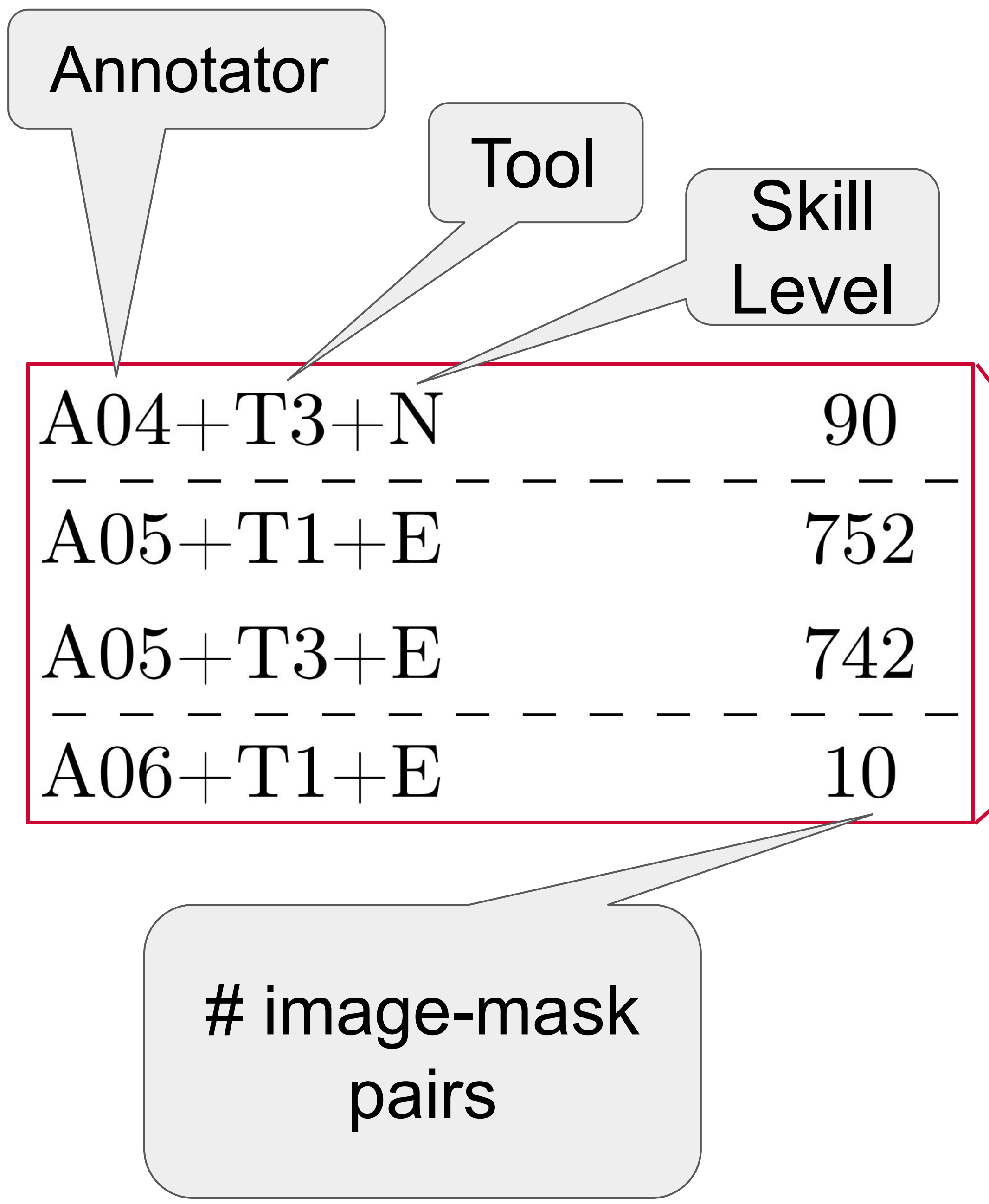




Quantitative Results on ISIC-MultiAnnot

Annotator + Tool + Experience	Seg. Count	1-StyleSeg	2-StyleSeg			3-StyleSeg			4-StyleSeg		
		$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$
A00+T2+E	1573	$0.892_{0.089}$	$0.923_{0.061}$	$0.913_{0.087}$	2	$0.944_{0.049}$	$0.913_{0.106}$	3	$0.944_{0.044}$	$0.914_{0.111}$	1
A00+T2+N	1305	$0.716_{0.302}$	$0.761_{0.293}$	$0.728_{0.308}$	2	$0.793_{0.287}$	$0.727_{0.313}$	3	$0.790_{0.290}$	$0.726_{0.304}$	3
A01+T1+N	6	$0.559_{0.362}$	$0.766_{0.152}$	$0.766_{0.152}$	1	$0.754_{0.132}$	$0.741_{0.125}$	2	$0.819_{0.106}$	$0.767_{0.113}$	2
A01+T3+E	297	$0.900_{0.104}$	$0.915_{0.093}$	$0.897_{0.107}$	2	$0.927_{0.075}$	$0.900_{0.097}$	1	$0.931_{0.067}$	$0.904_{0.090}$	3
A01+T3+N	2148	$0.829_{0.185}$	$0.857_{0.167}$	$0.817_{0.170}$	1	$0.869_{0.159}$	$0.836_{0.178}$	1	$0.876_{0.148}$	$0.836_{0.175}$	3
A02+T1+E	1742	$0.844_{0.177}$	$0.880_{0.140}$	$0.856_{0.159}$	1	$0.886_{0.132}$	$0.854_{0.159}$	1	$0.895_{0.112}$	$0.859_{0.148}$	4
A02+T3+E	468	$0.856_{0.172}$	$0.889_{0.167}$	$0.883_{0.175}$	2	$0.899_{0.161}$	$0.874_{0.188}$	3	$0.903_{0.146}$	$0.890_{0.160}$	1
A03+T1+E	1622	$0.778_{0.168}$	$0.845_{0.117}$	$0.827_{0.137}$	1	$0.854_{0.111}$	$0.824_{0.145}$	2	$0.881_{0.095}$	$0.823_{0.132}$	4
A03+T3+E	260	$0.891_{0.116}$	$0.912_{0.086}$	$0.876_{0.173}$	2	$0.923_{0.089}$	$0.868_{0.150}$	1	$0.932_{0.074}$	$0.874_{0.163}$	3
A04+T1+E	992	$0.850_{0.158}$	$0.880_{0.131}$	$0.860_{0.149}$	1	$0.888_{0.132}$	$0.866_{0.153}$	2	$0.906_{0.108}$	$0.856_{0.157}$	4
A04+T1+N	61	$0.760_{0.242}$	$0.840_{0.152}$	$0.823_{0.164}$	1	$0.837_{0.162}$	$0.786_{0.201}$	1	$0.827_{0.206}$	$0.789_{0.226}$	4
A04+T3+E	913	$0.912_{0.088}$	$0.939_{0.054}$	$0.934_{0.065}$	2	$0.948_{0.047}$	$0.926_{0.069}$	1	$0.951_{0.045}$	$0.932_{0.063}$	3
A04+T3+N	90	$0.877_{0.096}$	$0.910_{0.068}$	$0.905_{0.070}$	2	$0.928_{0.031}$	$0.908_{0.044}$	3	$0.926_{0.052}$	$0.913_{0.055}$	1
A05+T1+E	752	$0.815_{0.203}$	$0.862_{0.163}$	$0.837_{0.179}$	1	$0.873_{0.162}$	$0.827_{0.184}$	1	$0.882_{0.147}$	$0.841_{0.177}$	4
A05+T3+E	742	$0.875_{0.129}$	$0.903_{0.109}$	$0.891_{0.113}$	2	$0.916_{0.098}$	$0.878_{0.120}$	1	$0.919_{0.091}$	$0.891_{0.108}$	1
A06+T1+E	10	$0.824_{0.187}$	$0.902_{0.037}$	$0.885_{0.070}$	1	$0.909_{0.034}$	$0.889_{0.049}$	2	$0.909_{0.039}$	$0.880_{0.063}$	4
A06+T3+E	24	$0.862_{0.079}$	$0.916_{0.053}$	$0.916_{0.053}$	2	$0.934_{0.031}$	$0.923_{0.031}$	3	$0.933_{0.041}$	$0.929_{0.040}$	1
A07+T1+E	67	$0.820_{0.157}$	$0.877_{0.124}$	$0.867_{0.150}$	1	$0.890_{0.108}$	$0.862_{0.157}$	2	$0.897_{0.104}$	$0.862_{0.149}$	4
A07+T1+N	251	$0.837_{0.141}$	$0.892_{0.085}$	$0.879_{0.104}$	1	$0.903_{0.067}$	$0.875_{0.114}$	2	$0.905_{0.070}$	$0.873_{0.101}$	4
A07+T3+E	12	$0.925_{0.055}$	$0.938_{0.019}$	$0.937_{0.019}$	2	$0.939_{0.020}$	$0.916_{0.055}$	1	$0.947_{0.016}$	$0.932_{0.017}$	1
A07+T3+N	39	$0.863_{0.177}$	$0.918_{0.061}$	$0.913_{0.071}$	2	$0.933_{0.037}$	$0.899_{0.148}$	3	$0.934_{0.039}$	$0.914_{0.079}$	1
A08+T1+E	26	$0.666_{0.225}$	$0.750_{0.161}$	$0.680_{0.242}$	2	$0.747_{0.197}$	$0.653_{0.260}$	1	$0.793_{0.134}$	$0.666_{0.261}$	1
A08+T3+E	111	$0.605_{0.230}$	$0.668_{0.197}$	$0.626_{0.210}$	1	$0.677_{0.206}$	$0.628_{0.218}$	2	$0.735_{0.166}$	$0.669_{0.203}$	2
A09+T1+E	30	$0.815_{0.121}$	$0.841_{0.098}$	$0.784_{0.156}$	1	$0.873_{0.089}$	$0.833_{0.113}$	2	$0.884_{0.076}$	$0.812_{0.119}$	4
A09+T1+N	1	$0.953_{0.000}$	$0.927_{0.000}$	$0.927_{0.000}$	2	$0.955_{0.000}$	$0.955_{0.000}$	1	$0.947_{0.000}$	$0.947_{0.000}$	3
A09+T3+E	10	$0.900_{0.074}$	$0.918_{0.054}$	$0.918_{0.054}$	2	$0.933_{0.038}$	$0.909_{0.044}$	1	$0.937_{0.043}$	$0.919_{0.040}$	3
A09+T3+N	3	$0.894_{0.070}$	$0.911_{0.058}$	$0.911_{0.058}$	2	$0.957_{0.015}$	$0.957_{0.015}$	3	$0.944_{0.030}$	$0.944_{0.030}$	1

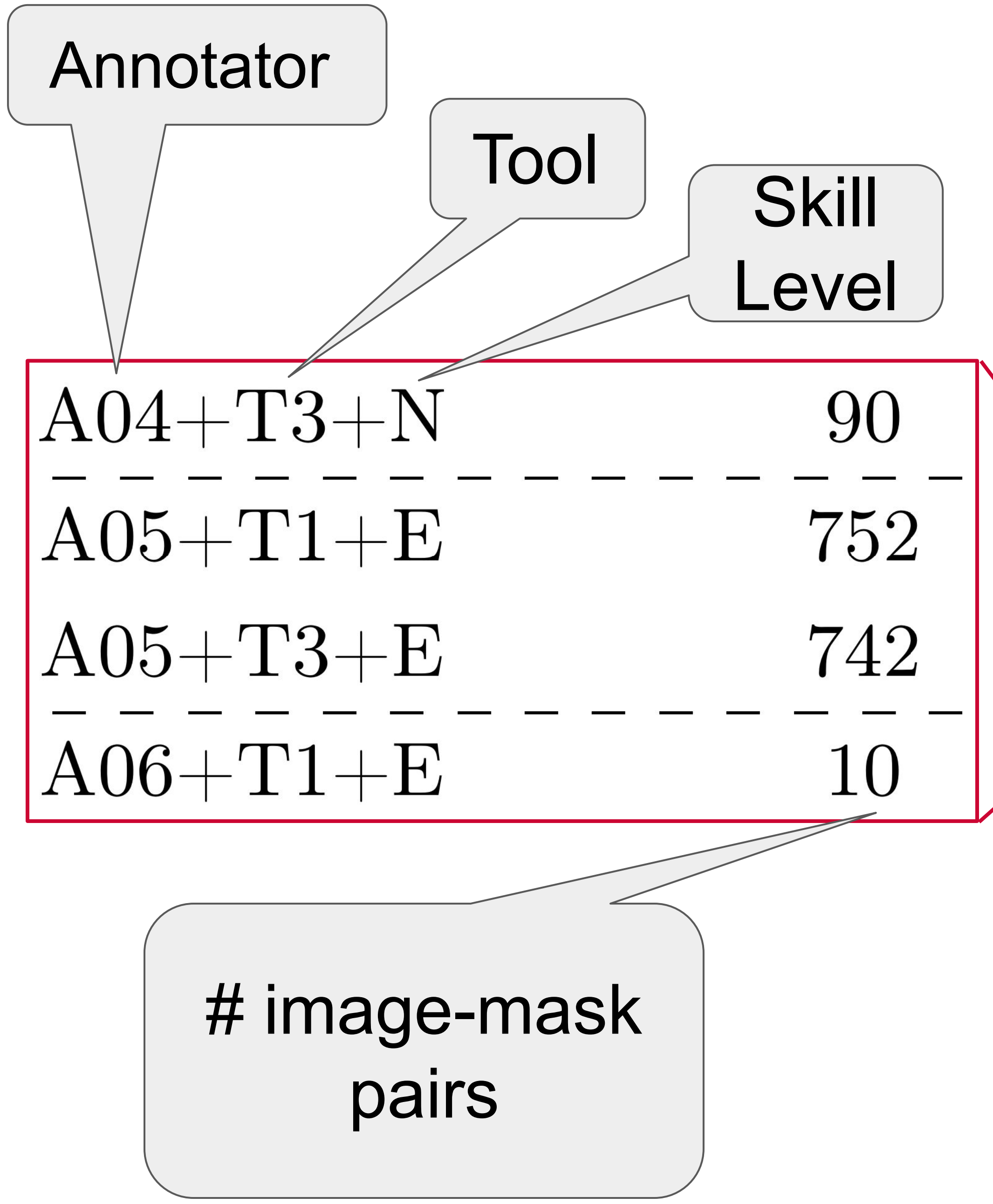
Quantitative Results on ISIC-MultiAnnot



Annotator + Tool + Experience	Seg. Count	1-StyleSeg	2-StyleSeg			3-StyleSeg			4-StyleSeg		
		$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$
A00+T2+E	1573	$0.892_{0.089}$	$0.923_{0.061}$	$0.913_{0.087}$	2	$0.944_{0.049}$	$0.913_{0.106}$	3	$0.944_{0.044}$	$0.914_{0.111}$	1
A00+T2+N	1305	$0.716_{0.302}$	$0.761_{0.293}$	$0.728_{0.308}$	2	$0.793_{0.287}$	$0.727_{0.313}$	3	$0.790_{0.290}$	$0.726_{0.304}$	3
A01+T1+N	6	$0.559_{0.362}$	$0.766_{0.152}$	$0.766_{0.152}$	1	$0.754_{0.132}$	$0.741_{0.125}$	2	$0.819_{0.106}$	$0.767_{0.113}$	2
A01+T3+E	297	$0.900_{0.104}$	$0.915_{0.093}$	$0.897_{0.107}$	2	$0.927_{0.075}$	$0.900_{0.097}$	1	$0.931_{0.067}$	$0.904_{0.090}$	3
A01+T3+N	2148	$0.829_{0.185}$	$0.857_{0.167}$	$0.817_{0.170}$	1	$0.869_{0.159}$	$0.836_{0.178}$	1	$0.876_{0.148}$	$0.836_{0.175}$	3
A02+T1+E	1742	$0.844_{0.177}$	$0.880_{0.140}$	$0.856_{0.159}$	1	$0.886_{0.132}$	$0.854_{0.159}$	1	$0.895_{0.112}$	$0.859_{0.148}$	4
A02+T3+E	468	$0.856_{0.172}$	$0.889_{0.167}$	$0.883_{0.175}$	2	$0.899_{0.161}$	$0.874_{0.188}$	3	$0.903_{0.146}$	$0.890_{0.160}$	1
A03+T1+E	1622	$0.778_{0.168}$	$0.845_{0.117}$	$0.827_{0.137}$	1	$0.854_{0.111}$	$0.824_{0.145}$	2	$0.881_{0.095}$	$0.823_{0.132}$	4
A03+T3+E	260	$0.891_{0.116}$	$0.912_{0.086}$	$0.876_{0.173}$	2	$0.923_{0.089}$	$0.868_{0.150}$	1	$0.932_{0.074}$	$0.874_{0.163}$	3
A04+T1+E	992	$0.850_{0.158}$	$0.880_{0.131}$	$0.860_{0.149}$	1	$0.888_{0.132}$	$0.866_{0.153}$	2	$0.906_{0.108}$	$0.856_{0.157}$	4
A04+T1+N	61	$0.760_{0.242}$	$0.840_{0.152}$	$0.823_{0.164}$	1	$0.837_{0.162}$	$0.786_{0.201}$	1	$0.827_{0.206}$	$0.789_{0.226}$	4
A04+T3+E	913	$0.912_{0.088}$	$0.939_{0.054}$	$0.934_{0.065}$	2	$0.948_{0.047}$	$0.926_{0.069}$	1	$0.951_{0.045}$	$0.932_{0.063}$	3
A04+T3+N	90	$0.877_{0.096}$	$0.910_{0.068}$	$0.905_{0.070}$	2	$0.928_{0.031}$	$0.908_{0.044}$	3	$0.926_{0.052}$	$0.913_{0.055}$	1
A05+T1+E	752	$0.815_{0.203}$	$0.862_{0.163}$	$0.837_{0.179}$	1	$0.873_{0.162}$	$0.827_{0.184}$	1	$0.882_{0.147}$	$0.841_{0.177}$	4
A05+T3+E	742	$0.875_{0.129}$	$0.903_{0.109}$	$0.891_{0.113}$	2	$0.916_{0.098}$	$0.878_{0.120}$	1	$0.919_{0.091}$	$0.891_{0.108}$	1
A06+T1+E	10	$0.824_{0.187}$	$0.902_{0.037}$	$0.885_{0.070}$	1	$0.909_{0.034}$	$0.889_{0.049}$	2	$0.909_{0.039}$	$0.880_{0.063}$	4
A06+T3+E	24	$0.862_{0.079}$	$0.916_{0.053}$	$0.916_{0.053}$	2	$0.934_{0.031}$	$0.923_{0.031}$	3	$0.933_{0.041}$	$0.929_{0.040}$	1
A07+T1+E	67	$0.820_{0.157}$	$0.877_{0.124}$	$0.867_{0.150}$	1	$0.890_{0.108}$	$0.862_{0.157}$	2	$0.897_{0.104}$	$0.862_{0.149}$	4
A07+T1+N	251	$0.837_{0.141}$	$0.892_{0.085}$	$0.879_{0.104}$	1	$0.903_{0.067}$	$0.875_{0.114}$	2	$0.905_{0.070}$	$0.873_{0.101}$	4
A07+T3+E	12	$0.925_{0.055}$	$0.938_{0.019}$	$0.937_{0.019}$	2	$0.939_{0.020}$	$0.916_{0.055}$	1	$0.947_{0.016}$	$0.932_{0.017}$	1
A07+T3+N	39	$0.863_{0.177}$	$0.918_{0.061}$	$0.913_{0.071}$	2	$0.933_{0.037}$	$0.899_{0.148}$	3	$0.934_{0.039}$	$0.914_{0.079}$	1
A08+T1+E	26	$0.666_{0.225}$	$0.750_{0.161}$	$0.680_{0.242}$	2	$0.747_{0.197}$	$0.653_{0.260}$	1	$0.793_{0.134}$	$0.666_{0.261}$	1
A08+T3+E	111	$0.605_{0.230}$	$0.668_{0.197}$	$0.626_{0.210}$	1	$0.677_{0.206}$	$0.628_{0.218}$	2	$0.735_{0.166}$	$0.669_{0.203}$	2
A09+T1+E	30	$0.815_{0.121}$	$0.841_{0.098}$	$0.784_{0.156}$	1	$0.873_{0.089}$	$0.833_{0.113}$	2	$0.884_{0.076}$	$0.812_{0.119}$	4
A09+T1+N	1	$0.953_{0.000}$	$0.927_{0.000}$	$0.927_{0.000}$	2	$0.955_{0.000}$	$0.955_{0.000}$	1	$0.947_{0.000}$	$0.947_{0.000}$	3
A09+T3+E	10	$0.900_{0.074}$	$0.918_{0.054}$	$0.918_{0.054}$	2	$0.933_{0.038}$	$0.909_{0.044}$	1	$0.937_{0.043}$	$0.919_{0.040}$	3
A09+T3+N	3	$0.894_{0.070}$	$0.911_{0.058}$	$0.911_{0.058}$	2	$0.957_{0.015}$	$0.957_{0.015}$	3	$0.944_{0.030}$	$0.944_{0.030}$	1

Quantitative Results on ISIC-MultiAnnot

4-StyleSeg



Annotator + Tool + Experience	Seg. Count	1-StyleSeg	2-StyleSeg			3-StyleSeg			4-StyleSeg		
		$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$
A00+T2+E	1573	$0.892_{0.089}$	$0.923_{0.061}$	$0.913_{0.087}$	2	$0.944_{0.049}$	$0.913_{0.106}$	3	$0.944_{0.044}$	$0.914_{0.111}$	1
A00+T2+N	1305	$0.716_{0.302}$	$0.761_{0.293}$	$0.728_{0.308}$	2	$0.793_{0.287}$	$0.727_{0.313}$	3	$0.790_{0.290}$	$0.726_{0.304}$	3
A01+T1+N	6	$0.559_{0.362}$	$0.766_{0.152}$	$0.766_{0.152}$	1	$0.754_{0.132}$	$0.741_{0.125}$	2	$0.819_{0.106}$	$0.767_{0.113}$	2
A01+T3+E	297	$0.900_{0.104}$	$0.915_{0.093}$	$0.897_{0.107}$	2	$0.927_{0.075}$	$0.900_{0.097}$	1	$0.931_{0.067}$	$0.904_{0.090}$	3
A01+T3+N	2148	$0.829_{0.185}$	$0.857_{0.167}$	$0.817_{0.170}$	1	$0.869_{0.159}$	$0.836_{0.178}$	1	$0.876_{0.148}$	$0.836_{0.175}$	3
A02+T1+E	1742	$0.844_{0.177}$	$0.880_{0.140}$	$0.856_{0.159}$	1	$0.886_{0.132}$	$0.854_{0.159}$	1	$0.895_{0.112}$	$0.859_{0.148}$	4
A02+T3+E	468	$0.856_{0.172}$	$0.889_{0.167}$	$0.883_{0.175}$	2	$0.899_{0.161}$	$0.874_{0.188}$	3	$0.903_{0.146}$	$0.890_{0.160}$	1
A03+T1+E	1622	$0.778_{0.168}$	$0.845_{0.117}$	$0.827_{0.137}$	1	$0.854_{0.111}$	$0.824_{0.145}$	2	$0.881_{0.095}$	$0.823_{0.132}$	4
A03+T3+E	260	$0.891_{0.116}$	$0.912_{0.086}$	$0.876_{0.173}$	2	$0.923_{0.089}$	$0.868_{0.150}$	1	$0.932_{0.074}$	$0.874_{0.163}$	3
A04+T1+E	992	$0.850_{0.158}$	$0.880_{0.131}$	$0.860_{0.149}$	1	$0.888_{0.132}$	$0.866_{0.153}$	2	$0.906_{0.108}$	$0.856_{0.157}$	4
A04+T1+N	61	$0.760_{0.242}$	$0.840_{0.152}$	$0.823_{0.164}$	1	$0.837_{0.162}$	$0.786_{0.201}$	1	$0.827_{0.206}$	$0.789_{0.226}$	4
A04+T3+E	913	$0.912_{0.088}$	$0.939_{0.054}$	$0.934_{0.065}$	2	$0.948_{0.047}$	$0.926_{0.069}$	1	$0.951_{0.045}$	$0.932_{0.063}$	3
A04+T3+N	90	$0.877_{0.096}$	$0.910_{0.068}$	$0.905_{0.070}$	2	$0.928_{0.031}$	$0.908_{0.044}$	3	$0.926_{0.052}$	$0.913_{0.055}$	1
A05+T1+E	752	$0.815_{0.203}$	$0.862_{0.163}$	$0.837_{0.179}$	1	$0.873_{0.162}$	$0.827_{0.184}$	1	$0.882_{0.147}$	$0.841_{0.177}$	4
A05+T3+E	742	$0.875_{0.129}$	$0.903_{0.109}$	$0.891_{0.113}$	2	$0.916_{0.098}$	$0.878_{0.120}$	1	$0.919_{0.091}$	$0.891_{0.108}$	1
A06+T1+E	10	$0.824_{0.187}$	$0.902_{0.037}$	$0.885_{0.070}$	1	$0.909_{0.034}$	$0.889_{0.049}$	2	$0.909_{0.039}$	$0.880_{0.063}$	4
A06+T3+E	24	$0.862_{0.079}$	$0.916_{0.053}$	$0.916_{0.053}$	2	$0.934_{0.031}$	$0.923_{0.031}$	3	$0.933_{0.041}$	$0.929_{0.040}$	1
A07+T1+E	67	$0.820_{0.157}$	$0.877_{0.124}$	$0.867_{0.150}$	1	$0.890_{0.108}$	$0.862_{0.157}$	2	$0.897_{0.104}$	$0.862_{0.149}$	4
A07+T1+N	251	$0.837_{0.141}$	$0.892_{0.085}$	$0.879_{0.104}$	1	$0.903_{0.067}$	$0.875_{0.114}$	2	$0.905_{0.070}$	$0.873_{0.101}$	4
A07+T3+E	12	$0.925_{0.055}$	$0.938_{0.019}$	$0.937_{0.019}$	2	$0.939_{0.020}$	$0.916_{0.055}$	1	$0.947_{0.016}$	$0.932_{0.017}$	1
A07+T3+N	39	$0.863_{0.177}$	$0.918_{0.061}$	$0.913_{0.071}$	2	$0.933_{0.037}$	$0.899_{0.148}$	3	$0.934_{0.039}$	$0.914_{0.079}$	1
A08+T1+E	26	$0.666_{0.225}$	$0.750_{0.161}$	$0.680_{0.242}$	2	$0.747_{0.197}$	$0.653_{0.260}$	1	$0.793_{0.134}$	$0.666_{0.261}$	1
A08+T3+E	111	$0.605_{0.230}$	$0.668_{0.197}$	$0.626_{0.210}$	1	$0.677_{0.206}$	$0.628_{0.218}$	2	$0.735_{0.166}$	$0.669_{0.203}$	2
A09+T1+E	30	$0.815_{0.121}$	$0.841_{0.098}$	$0.784_{0.156}$	1	$0.873_{0.089}$	$0.833_{0.113}$	2	$0.884_{0.076}$	$0.812_{0.119}$	4
A09+T1+N	1	$0.953_{0.000}$	$0.927_{0.000}$	$0.927_{0.000}$	2	$0.955_{0.000}$	$0.955_{0.000}$	1	$0.947_{0.000}$	$0.947_{0.000}$	3
A09+T3+E	10	$0.900_{0.074}$	$0.918_{0.054}$	$0.918_{0.054}$	2	$0.933_{0.038}$	$0.909_{0.044}$	1	$0.937_{0.043}$	$0.919_{0.040}$	3
A09+T3+N	3	$0.894_{0.070}$	$0.911_{0.058}$	$0.911_{0.058}$	2	$0.957_{0.015}$	$0.957_{0.015}$	3	$0.944_{0.030}$	$0.944_{0.030}$	1

Annotator + Tool + Experience	Seg. Count	1-StyleSeg	2-StyleSeg			3-StyleSeg			4-StyleSeg		
		$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$
A00+T2+E	1573	$0.892_{0.089}$	$0.923_{0.061}$	$0.913_{0.087}$	2	$0.944_{0.049}$	$0.913_{0.106}$	3	$0.944_{0.044}$	$0.914_{0.111}$	1
A00+T2+N	1305	$0.716_{0.302}$	$0.761_{0.293}$	$0.728_{0.308}$	2	$0.793_{0.287}$	$0.727_{0.313}$	3	$0.790_{0.290}$	$0.726_{0.304}$	3
A01+T1+N	6	$0.559_{0.362}$	$0.766_{0.152}$	$0.766_{0.152}$	1	$0.754_{0.132}$	$0.741_{0.125}$	2	$0.819_{0.106}$	$0.767_{0.113}$	2
A01+T3+E	297	$0.900_{0.104}$	$0.915_{0.093}$	$0.897_{0.107}$	2	$0.927_{0.075}$	$0.900_{0.097}$	1	$0.931_{0.067}$	$0.904_{0.090}$	3
A01+T3+N	2148	$0.829_{0.185}$	$0.857_{0.167}$	$0.817_{0.170}$	1	$0.869_{0.159}$	$0.836_{0.178}$	1	$0.876_{0.148}$	$0.836_{0.175}$	3
A02+T1+E	1742	$0.844_{0.177}$	$0.880_{0.140}$	$0.856_{0.159}$	1	$0.886_{0.132}$	$0.854_{0.159}$	1	$0.895_{0.112}$	$0.859_{0.148}$	4
A02+T3+E	468	$0.856_{0.172}$	$0.889_{0.167}$	$0.883_{0.175}$	2	$0.899_{0.161}$	$0.874_{0.188}$	3	$0.903_{0.146}$	$0.890_{0.160}$	1
A03+T1+E	1622	$0.778_{0.168}$	$0.845_{0.117}$	$0.827_{0.137}$	1	$0.854_{0.111}$	$0.824_{0.145}$	2	$0.881_{0.095}$	$0.823_{0.132}$	4
A03+T3+E	260	$0.891_{0.116}$	$0.912_{0.086}$	$0.876_{0.173}$	2	$0.923_{0.089}$	$0.868_{0.150}$	1	$0.932_{0.074}$	$0.874_{0.163}$	3
A04+T1+E	992	$0.850_{0.158}$	$0.880_{0.131}$	$0.860_{0.149}$	1	$0.888_{0.132}$	$0.866_{0.153}$	2	$0.906_{0.108}$	$0.856_{0.157}$	4
A04+T1+N	61	$0.760_{0.242}$	$0.840_{0.152}$	$0.823_{0.164}$	1	$0.837_{0.162}$	$0.786_{0.201}$	1	$0.827_{0.206}$	$0.789_{0.226}$	4
A04+T3+E	913	$0.912_{0.088}$	$0.939_{0.054}$	$0.934_{0.065}$	2	$0.948_{0.047}$	$0.926_{0.069}$	1	$0.951_{0.045}$	$0.932_{0.063}$	3
A04+T3+N	90	$0.877_{0.096}$	$0.910_{0.068}$	$0.905_{0.070}$	2	$0.928_{0.031}$	$0.908_{0.044}$	3	$0.926_{0.052}$	$0.913_{0.055}$	1
A05+T1+E	752	$0.815_{0.203}$	$0.862_{0.163}$	$0.837_{0.179}$	1	$0.873_{0.162}$	$0.827_{0.184}$	1	$0.882_{0.147}$	$0.841_{0.177}$	4
A05+T3+E	742	$0.875_{0.129}$	$0.903_{0.109}$	$0.891_{0.113}$	2	$0.916_{0.098}$	$0.878_{0.120}$	1	$0.919_{0.091}$	$0.891_{0.108}$	1
A06+T1+E	10	$0.824_{0.187}$	$0.902_{0.037}$	$0.885_{0.070}$	1	$0.909_{0.034}$	$0.889_{0.049}$	2	$0.909_{0.039}$	$0.880_{0.063}$	4
A06+T3+E	24	$0.862_{0.079}$	$0.916_{0.053}$	$0.916_{0.053}$	2	$0.934_{0.031}$	$0.923_{0.031}$	3	$0.933_{0.041}$	$0.929_{0.040}$	1
A07+T1+E	67	$0.820_{0.157}$	$0.877_{0.124}$	$0.867_{0.150}$	1	$0.890_{0.108}$	$0.862_{0.157}$	2	$0.897_{0.104}$	$0.862_{0.149}$	4
A07+T1+N	251	$0.837_{0.141}$	$0.892_{0.085}$	$0.879_{0.104}$	1	$0.903_{0.067}$	$0.875_{0.114}$	2	$0.905_{0.070}$	$0.873_{0.101}$	4
A07+T3+E	12	$0.925_{0.055}$	$0.938_{0.019}$	$0.937_{0.019}$	2	$0.939_{0.020}$	$0.916_{0.055}$	1	$0.947_{0.016}$	$0.932_{0.017}$	1
A07+T3+N	39	$0.863_{0.177}$	$0.918_{0.061}$	$0.913_{0.071}$	2	$0.933_{0.037}$	$0.899_{0.148}$	3	$0.934_{0.039}$	$0.914_{0.079}$	1
A08+T1+E	26	$0.666_{0.225}$	$0.750_{0.161}$	$0.680_{0.242}$	2	$0.747_{0.197}$	$0.653_{0.260}$	1	$0.793_{0.134}$	$0.666_{0.261}$	1
A08+T3+E	111	$0.605_{0.230}$	$0.668_{0.197}$	$0.626_{0.210}$	1	$0.677_{0.206}$	$0.628_{0.218}$	2	$0.735_{0.166}$	$0.669_{0.203}$	2
A09+T1+E	30	$0.815_{0.121}$	$0.841_{0.098}$	$0.784_{0.156}$	1	$0.873_{0.089}$	$0.833_{0.113}$	2	$0.884_{0.076}$	$0.812_{0.119}$	4
A09+T1+N	1	$0.953_{0.000}$	$0.927_{0.000}$	$0.927_{0.000}$	2	$0.955_{0.000}$	$0.955_{0.000}$	1	$0.947_{0.000}$	$0.947_{0.000}$	3
A09+T3+E	10	$0.900_{0.074}$	$0.918_{0.054}$	$0.918_{0.054}$	2	$0.933_{0.038}$	$0.909_{0.044}$	1	$0.937_{0.043}$	$0.919_{0.040}$	3
A09+T3+N	3	$0.894_{0.070}$	$0.911_{0.058}$	$0.911_{0.058}$	2	$0.957_{0.015}$	$0.957_{0.015}$	3	$0.944_{0.030}$	$0.944_{0.030}$	1

ISIC-MultiAnnot Results: Key Takeaways

1. **Improved diversity without compromising quality**: for all *M* ≥ 2, choosing a single style that, for each annotator preference, maximizes agreement with the "ground truth" still outperforms 1-StyleSeg.

ISIC-MultiAnnot Results: Key Takeaways

Annotator + Tool + Experience	Seg. Count	1-StyleSeg	2-StyleSeg			3-StyleSeg			4-StyleSeg		
		$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$
A00+T2+E	1573	$0.892_{0.089}$	$0.923_{0.061}$	$0.913_{0.087}$	2	$0.944_{0.049}$	$0.913_{0.106}$	3	$0.944_{0.044}$	$0.914_{0.111}$	1
A00+T2+N	1305	$0.716_{0.302}$	$0.761_{0.293}$	$0.728_{0.308}$	2	$0.793_{0.287}$	$0.727_{0.313}$	3	$0.790_{0.290}$	$0.726_{0.304}$	3
A01+T1+N	6	$0.559_{0.362}$	$0.766_{0.152}$	$0.766_{0.152}$	1	$0.754_{0.132}$	$0.741_{0.125}$	2	$0.819_{0.106}$	$0.767_{0.113}$	2
A01+T3+E	297	$0.900_{0.104}$	$0.915_{0.093}$	$0.897_{0.107}$	2	$0.927_{0.075}$	$0.900_{0.097}$	1	$0.931_{0.067}$	$0.904_{0.090}$	3
A01+T3+N	2148	$0.829_{0.185}$	$0.857_{0.167}$	$0.817_{0.170}$	1	$0.869_{0.159}$	$0.836_{0.178}$	1	$0.876_{0.148}$	$0.836_{0.175}$	3
A02+T1+E	1742	$0.844_{0.177}$	$0.880_{0.140}$	$0.856_{0.159}$	1	$0.886_{0.132}$	$0.854_{0.159}$	1	$0.895_{0.112}$	$0.859_{0.148}$	4
A02+T3+E	468	$0.856_{0.172}$	$0.889_{0.167}$	$0.883_{0.175}$	2	$0.899_{0.161}$	$0.874_{0.188}$	3	$0.903_{0.146}$	$0.890_{0.160}$	1
A03+T1+E	1622	$0.778_{0.168}$	$0.845_{0.117}$	$0.827_{0.137}$	1	$0.854_{0.111}$	$0.824_{0.145}$	2	$0.881_{0.095}$	$0.823_{0.132}$	4
A03+T3+E	260	$0.891_{0.116}$	$0.912_{0.086}$	$0.876_{0.173}$	2	$0.923_{0.089}$	$0.868_{0.150}$	1	$0.932_{0.074}$	$0.874_{0.163}$	3
A04+T1+E	992	$0.850_{0.158}$	$0.880_{0.131}$	$0.860_{0.149}$	1	$0.888_{0.132}$	$0.866_{0.153}$	2	$0.906_{0.108}$	$0.856_{0.157}$	4
A04+T1+N	61	$0.760_{0.242}$	$0.840_{0.152}$	$0.823_{0.164}$	1	$0.837_{0.162}$	$0.786_{0.201}$	1	$0.827_{0.206}$	$0.789_{0.226}$	4
A04+T3+E	913	$0.912_{0.088}$	$0.939_{0.054}$	$0.934_{0.065}$	2	$0.948_{0.047}$	$0.926_{0.069}$	1	$0.951_{0.045}$	$0.932_{0.063}$	3
A04+T3+N	90	$0.877_{0.096}$	$0.910_{0.068}$	$0.905_{0.070}$	2	$0.928_{0.031}$	$0.908_{0.044}$	3	$0.926_{0.052}$	$0.913_{0.055}$	1
A05+T1+E	752	$0.815_{0.203}$	$0.862_{0.163}$	$0.837_{0.179}$	1	$0.873_{0.162}$	$0.827_{0.184}$	1	$0.882_{0.147}$	$0.841_{0.177}$	4
A05+T3+E	742	$0.875_{0.129}$	$0.903_{0.109}$	$0.891_{0.113}$	2	$0.916_{0.098}$	$0.878_{0.120}$	1	$0.919_{0.091}$	$0.891_{0.108}$	1
A06+T1+E	10	$0.824_{0.187}$	$0.902_{0.037}$	$0.885_{0.070}$	1	$0.909_{0.034}$	$0.889_{0.049}$	2	$0.909_{0.039}$	$0.880_{0.063}$	4
A06+T3+E	24	$0.862_{0.079}$	$0.916_{0.053}$	$0.916_{0.053}$	2	$0.934_{0.031}$	$0.923_{0.031}$	3	$0.933_{0.041}$	$0.929_{0.040}$	1
A07+T1+E	67	$0.820_{0.157}$	$0.877_{0.124}$	$0.867_{0.150}$	1	$0.890_{0.108}$	$0.862_{0.157}$	2	$0.897_{0.104}$	$0.862_{0.149}$	4
A07+T1+N	251	$0.837_{0.141}$	$0.892_{0.085}$	$0.879_{0.104}$	1	$0.903_{0.067}$	$0.875_{0.114}$	2	$0.905_{0.070}$	$0.873_{0.101}$	4
A07+T3+E	12	$0.925_{0.055}$	$0.938_{0.019}$	$0.937_{0.019}$	2	$0.939_{0.020}$	$0.916_{0.055}$	1	$0.947_{0.016}$	$0.932_{0.017}$	1
A07+T3+N	39	$0.863_{0.177}$	$0.918_{0.061}$	$0.913_{0.071}$	2	$0.933_{0.037}$	$0.899_{0.148}$	3	$0.934_{0.039}$	$0.914_{0.079}$	1
A08+T1+E	26	$0.666_{0.225}$	$0.750_{0.161}$	$0.680_{0.242}$	2	$0.747_{0.197}$	$0.653_{0.260}$	1	$0.793_{0.134}$	$0.666_{0.261}$	1
A08+T3+E	111	$0.605_{0.230}$	$0.668_{0.197}$	$0.626_{0.210}$	1	$0.677_{0.206}$	$0.628_{0.218}$	2	$0.735_{0.166}$	$0.669_{0.203}$	2
A09+T1+E	30	$0.815_{0.121}$	$0.841_{0.098}$	$0.784_{0.156}$	1	$0.873_{0.089}$	$0.833_{0.113}$	2	$0.884_{0.076}$	$0.812_{0.119}$	4
A09+T1+N	1	$0.953_{0.000}$	$0.927_{0.000}$	$0.927_{0.000}$	2	$0.955_{0.000}$	$0.955_{0.000}$	1	$0.947_{0.000}$	$0.947_{0.000}$	3
A09+T3+E	10	$0.900_{0.074}$	$0.918_{0.054}$	$0.918_{0.054}$	2	$0.933_{0.038}$	$0.909_{0.044}$	1	$0.937_{0.043}$	$0.919_{0.040}$	3
A09+T3+N	3	$0.894_{0.070}$	$0.911_{0.058}$	$0.911_{0.058}$	2	$0.957_{0.015}$	$0.957_{0.015}$	3	$0.944_{0.030}$	$0.944_{0.030}$	1

1. **Improved diversity without compromising quality**: for all $M \geq 2$, choosing a single style that, for each annotator preference, maximizes agreement with the “ground truth” still outperforms 1-StyleSeg.

Personalization in segmentation: each user can choose their own style.

ISIC-MultiAnnot Results: Key Takeaways

Annotator + Tool + Experience	Seg. Count	1-StyleSeg	2-StyleSeg			3-StyleSeg			4-StyleSeg		
		$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$
A00+T2+E	1573	$0.892_{0.089}$	$0.923_{0.061}$	$0.913_{0.087}$	2	$0.944_{0.049}$	$0.913_{0.106}$	3	$0.944_{0.044}$	$0.914_{0.111}$	1
A00+T2+N	1305	$0.716_{0.302}$	$0.761_{0.293}$	$0.728_{0.308}$	2	$0.793_{0.287}$	$0.727_{0.313}$	3	$0.790_{0.290}$	$0.726_{0.304}$	3
A01+T1+N	6	$0.559_{0.362}$	$0.766_{0.152}$	$0.766_{0.152}$	1	$0.754_{0.132}$	$0.741_{0.125}$	2	$0.819_{0.106}$	$0.767_{0.113}$	2
A01+T3+E	297	$0.900_{0.104}$	$0.915_{0.093}$	$0.897_{0.107}$	2	$0.927_{0.075}$	$0.900_{0.097}$	1	$0.931_{0.067}$	$0.904_{0.090}$	3
A01+T3+N	2148	$0.829_{0.185}$	$0.857_{0.167}$	$0.817_{0.170}$	1	$0.869_{0.159}$	$0.836_{0.178}$	1	$0.876_{0.148}$	$0.836_{0.175}$	3
A02+T1+E	1742	$0.844_{0.177}$	$0.880_{0.140}$	$0.856_{0.159}$	1	$0.886_{0.132}$	$0.854_{0.159}$	1	$0.895_{0.112}$	$0.859_{0.148}$	4
A02+T3+E	468	$0.856_{0.172}$	$0.889_{0.167}$	$0.883_{0.175}$	2	$0.899_{0.161}$	$0.874_{0.188}$	3	$0.903_{0.146}$	$0.890_{0.160}$	1
A03+T1+E	1622	$0.778_{0.168}$	$0.845_{0.117}$	$0.827_{0.137}$	1	$0.854_{0.111}$	$0.824_{0.145}$	2	$0.881_{0.095}$	$0.823_{0.132}$	4
A03+T3+E	260	$0.891_{0.116}$	$0.912_{0.086}$	$0.876_{0.173}$	2	$0.923_{0.089}$	$0.868_{0.150}$	1	$0.932_{0.074}$	$0.874_{0.163}$	3
A04+T1+E	992	$0.850_{0.158}$	$0.880_{0.131}$	$0.860_{0.149}$	1	$0.888_{0.132}$	$0.866_{0.153}$	2	$0.906_{0.108}$	$0.856_{0.157}$	4
A04+T1+N	61	$0.760_{0.242}$	$0.840_{0.152}$	$0.823_{0.164}$	1	$0.837_{0.162}$	$0.786_{0.201}$	1	$0.827_{0.206}$	$0.789_{0.226}$	4
A04+T3+E	913	$0.912_{0.088}$	$0.939_{0.054}$	$0.934_{0.065}$	2	$0.948_{0.047}$	$0.926_{0.069}$	1	$0.951_{0.045}$	$0.932_{0.063}$	3
A04+T3+N	90	$0.877_{0.096}$	$0.910_{0.068}$	$0.905_{0.070}$	2	$0.928_{0.031}$	$0.908_{0.044}$	3	$0.926_{0.052}$	$0.913_{0.055}$	1
A05+T1+E	752	$0.815_{0.203}$	$0.862_{0.163}$	$0.837_{0.179}$	1	$0.873_{0.162}$	$0.827_{0.184}$	1	$0.882_{0.147}$	$0.841_{0.177}$	4
A05+T3+E	742	$0.875_{0.129}$	$0.903_{0.109}$	$0.891_{0.113}$	2	$0.916_{0.098}$	$0.878_{0.120}$	1	$0.919_{0.091}$	$0.891_{0.108}$	1
A06+T1+E	10	$0.824_{0.187}$	$0.902_{0.037}$	$0.885_{0.070}$	1	$0.909_{0.034}$	$0.889_{0.049}$	2	$0.909_{0.039}$	$0.880_{0.063}$	4
A06+T3+E	24	$0.862_{0.079}$	$0.916_{0.053}$	$0.916_{0.053}$	2	$0.934_{0.031}$	$0.923_{0.031}$	3	$0.933_{0.041}$	$0.929_{0.040}$	1
A07+T1+E	67	$0.820_{0.157}$	$0.877_{0.124}$	$0.867_{0.150}$	1	$0.890_{0.108}$	$0.862_{0.157}$	2	$0.897_{0.104}$	$0.862_{0.149}$	4
A07+T1+N	251	$0.837_{0.141}$	$0.892_{0.085}$	$0.879_{0.104}$	1	$0.903_{0.067}$	$0.875_{0.114}$	2	$0.905_{0.070}$	$0.873_{0.101}$	4
A07+T3+E	12	$0.925_{0.055}$	$0.938_{0.019}$	$0.937_{0.019}$	2	$0.939_{0.020}$	$0.916_{0.055}$	1	$0.947_{0.016}$	$0.932_{0.017}$	1
A07+T3+N	39	$0.863_{0.177}$	$0.918_{0.061}$	$0.913_{0.071}$	2	$0.933_{0.037}$	$0.899_{0.148}$	3	$0.934_{0.039}$	$0.914_{0.079}$	1
A08+T1+E	26	$0.666_{0.225}$	$0.750_{0.161}$	$0.680_{0.242}$	2	$0.747_{0.197}$	$0.653_{0.260}$	1	$0.793_{0.134}$	$0.666_{0.261}$	1
A08+T3+E	111	$0.605_{0.230}$	$0.668_{0.197}$	$0.626_{0.210}$	1	$0.677_{0.206}$	$0.628_{0.218}$	2	$0.735_{0.166}$	$0.669_{0.203}$	2
A09+T1+E	30	$0.815_{0.121}$	$0.841_{0.098}$	$0.784_{0.156}$	1	$0.873_{0.089}$	$0.833_{0.113}$	2	$0.884_{0.076}$	$0.812_{0.119}$	4
A09+T1+N	1	$0.953_{0.000}$	$0.927_{0.000}$	$0.927_{0.000}$	2	$0.955_{0.000}$	$0.955_{0.000}$	1	$0.947_{0.000}$	$0.947_{0.000}$	3
A09+T3+E	10	$0.900_{0.074}$	$0.918_{0.054}$	$0.918_{0.054}$	2	$0.933_{0.038}$	$0.909_{0.044}$	1	$0.937_{0.043}$	$0.919_{0.040}$	3
A09+T3+N	3	$0.894_{0.070}$	$0.911_{0.058}$	$0.911_{0.058}$	2	$0.957_{0.015}$	$0.957_{0.015}$	3	$0.944_{0.030}$	$0.944_{0.030}$	1

1. **Improved diversity without compromising quality**: for all *M* ≥ 2, choosing a single style that, for each annotator preference, maximizes agreement with the “ground truth” still outperforms 1-StyleSeg.

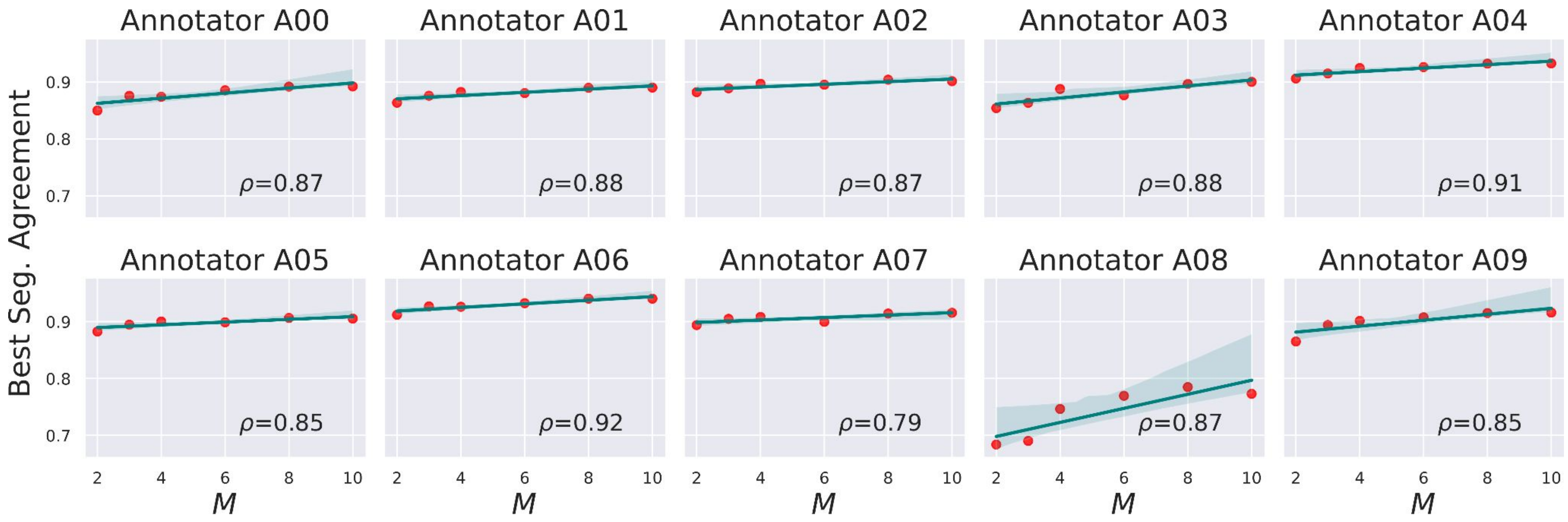
2. **Performance improves as *M* increases.**

Annotator + Tool + Experience	Seg. Count	1-StyleSeg	2-StyleSeg			3-StyleSeg			4-StyleSeg		
		$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$
A00+T2+E	1573	$0.892_{0.089}$	$0.923_{0.061}$	$0.913_{0.087}$	2	$0.944_{0.049}$	$0.913_{0.106}$	3	$0.944_{0.044}$	$0.914_{0.111}$	1
A00+T2+N	1305	$0.716_{0.302}$	$0.761_{0.293}$	$0.728_{0.308}$	2	$0.793_{0.287}$	$0.727_{0.313}$	3	$0.790_{0.290}$	$0.726_{0.304}$	3
A01+T1+N	6	$0.559_{0.362}$	$0.766_{0.152}$	$0.766_{0.152}$	1	$0.754_{0.132}$	$0.741_{0.125}$	2	$0.819_{0.106}$	$0.767_{0.113}$	2
A01+T3+E	297	$0.900_{0.104}$	$0.915_{0.093}$	$0.897_{0.107}$	2	$0.927_{0.075}$	$0.900_{0.097}$	1	$0.931_{0.067}$	$0.904_{0.090}$	3
A01+T3+N	2148	$0.829_{0.185}$	$0.857_{0.167}$	$0.817_{0.170}$	1	$0.869_{0.159}$	$0.836_{0.178}$	1	$0.876_{0.148}$	$0.836_{0.175}$	3
A02+T1+E	1742	$0.844_{0.177}$	$0.880_{0.140}$	$0.856_{0.159}$	1	$0.886_{0.132}$	$0.854_{0.159}$	1	$0.895_{0.112}$	$0.859_{0.148}$	4
A02+T3+E	468	$0.856_{0.172}$	$0.889_{0.167}$	$0.883_{0.175}$	2	$0.899_{0.161}$	$0.874_{0.188}$	3	$0.903_{0.146}$	$0.890_{0.160}$	1
A03+T1+E	1622	$0.778_{0.168}$	$0.845_{0.117}$	$0.827_{0.137}$	1	$0.854_{0.111}$	$0.824_{0.145}$	2	$0.881_{0.095}$	$0.823_{0.132}$	4
A03+T3+E	260	$0.891_{0.116}$	$0.912_{0.086}$	$0.876_{0.173}$	2	$0.923_{0.089}$	$0.868_{0.150}$	1	$0.932_{0.074}$	$0.874_{0.163}$	3
A04+T1+E	992	$0.850_{0.158}$	$0.880_{0.131}$	$0.860_{0.149}$	1	$0.888_{0.132}$	$0.866_{0.153}$	2	$0.906_{0.108}$	$0.856_{0.157}$	4
A04+T1+N	61	$0.760_{0.242}$	$0.840_{0.152}$	$0.823_{0.164}$	1	$0.837_{0.162}$	$0.786_{0.201}$	1	$0.827_{0.206}$	$0.789_{0.226}$	4
A04+T3+E	913	$0.912_{0.088}$	$0.939_{0.054}$	$0.934_{0.065}$	2	$0.948_{0.047}$	$0.926_{0.069}$	1	$0.951_{0.045}$	$0.932_{0.063}$	3
A04+T3+N	90	$0.877_{0.096}$	$0.910_{0.068}$	$0.905_{0.070}$	2	$0.928_{0.031}$	$0.908_{0.044}$	3	$0.926_{0.052}$	$0.913_{0.055}$	1
A05+T1+E	752	$0.815_{0.203}$	$0.862_{0.163}$	$0.837_{0.179}$	1	$0.873_{0.162}$	$0.827_{0.184}$	1	$0.882_{0.147}$	$0.841_{0.177}$	4
A05+T3+E	742	$0.875_{0.129}$	$0.903_{0.109}$	$0.891_{0.113}$	2	$0.916_{0.098}$	$0.878_{0.120}$	1	$0.919_{0.091}$	$0.891_{0.108}$	1
A06+T1+E	10	$0.824_{0.187}$	$0.902_{0.037}$	$0.885_{0.070}$	1	$0.909_{0.034}$	$0.889_{0.049}$	2	$0.909_{0.039}$	$0.880_{0.063}$	4
A06+T3+E	24	$0.862_{0.079}$	$0.916_{0.053}$	$0.916_{0.053}$	2	$0.934_{0.031}$	$0.923_{0.031}$	3	$0.933_{0.041}$	$0.929_{0.040}$	1
A07+T1+E	67	$0.820_{0.157}$	$0.877_{0.124}$	$0.867_{0.150}$	1	$0.890_{0.108}$	$0.862_{0.157}$	2	$0.897_{0.104}$	$0.862_{0.149}$	4
A07+T1+N	251	$0.837_{0.141}$	$0.892_{0.085}$	$0.879_{0.104}$	1	$0.903_{0.067}$	$0.875_{0.114}$	2	$0.905_{0.070}$	$0.873_{0.101}$	4
A07+T3+E	12	$0.925_{0.055}$	$0.938_{0.019}$	$0.937_{0.019}$	2	$0.939_{0.020}$	$0.916_{0.055}$	1	$0.947_{0.016}$	$0.932_{0.017}$	1
A07+T3+N	39	$0.863_{0.177}$	$0.918_{0.061}$	$0.913_{0.071}$	2	$0.933_{0.037}$	$0.899_{0.148}$	3	$0.934_{0.039}$	$0.914_{0.079}$	1
A08+T1+E	26	$0.666_{0.225}$	$0.750_{0.161}$	$0.680_{0.242}$	2	$0.747_{0.197}$	$0.653_{0.260}$	1	$0.793_{0.134}$	$0.666_{0.261}$	1
A08+T3+E	111	$0.605_{0.230}$	$0.668_{0.197}$	$0.626_{0.210}$	1	$0.677_{0.206}$	$0.628_{0.218}$	2	$0.735_{0.166}$	$0.669_{0.203}$	2
A09+T1+E	30	$0.815_{0.121}$	$0.841_{0.098}$	$0.784_{0.156}$	1	$0.873_{0.089}$	$0.833_{0.113}$	2	$0.884_{0.076}$	$0.812_{0.119}$	4
A09+T1+N	1	$0.953_{0.000}$	$0.927_{0.000}$	$0.927_{0.000}$	2	$0.955_{0.000}$	$0.955_{0.000}$	1	$0.947_{0.000}$	$0.947_{0.000}$	3
A09+T3+E	10	$0.900_{0.074}$	$0.918_{0.054}$	$0.918_{0.054}$	2	$0.933_{0.038}$	$0.909_{0.044}$	1	$0.937_{0.043}$	$0.919_{0.040}$	3
A09+T3+N	3	$0.894_{0.070}$	$0.911_{0.058}$	$0.911_{0.058}$	2	$0.957_{0.015}$	$0.957_{0.015}$	3	$0.944_{0.030}$	$0.944_{0.030}$	1

ISIC-MultiAnnot Results: Key Takeaways

1. **Improved diversity without compromising quality**: for all $M \geq 2$, choosing a single style that, for each annotator preference, maximizes agreement with the “ground truth” still outperforms 1-StyleSeg.

2. **Performance improves as *M* increases.**



ISIC-MultiAnnot Results: Key Takeaways

Annotator + Tool + Experience	Seg. Count	1-StyleSeg	2-StyleSeg			3-StyleSeg			4-StyleSeg		
		$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$	$\text{Dice}_{\text{ISSS}}$	$\text{Dice}_{\text{ASSS}}$	$\mathcal{J}$
A00+T2+E	1573	$0.892_{0.089}$	$0.923_{0.061}$	$0.913_{0.087}$	2	$0.944_{0.049}$	$0.913_{0.106}$	3	$0.944_{0.044}$	$0.914_{0.111}$	1
A00+T2+N	1305	$0.716_{0.302}$	$0.761_{0.293}$	$0.728_{0.308}$	2	$0.793_{0.287}$	$0.727_{0.313}$	3	$0.790_{0.290}$	$0.726_{0.304}$	3
A01+T1+N	6	$0.559_{0.362}$	$0.766_{0.152}$	$0.766_{0.152}$	1	$0.754_{0.132}$	$0.741_{0.125}$	2	$0.819_{0.106}$	$0.767_{0.113}$	2
A01+T3+E	297	$0.900_{0.104}$	$0.915_{0.093}$	$0.897_{0.107}$	2	$0.927_{0.075}$	$0.900_{0.097}$	1	$0.931_{0.067}$	$0.904_{0.090}$	3
A01+T3+N	2148	$0.829_{0.185}$	$0.857_{0.167}$	$0.817_{0.170}$	1	$0.869_{0.159}$	$0.836_{0.178}$	1	$0.876_{0.148}$	$0.836_{0.175}$	3
A02+T1+E	1742	$0.844_{0.177}$	$0.880_{0.140}$	$0.856_{0.159}$	1	$0.886_{0.132}$	$0.854_{0.159}$	1	$0.895_{0.112}$	$0.859_{0.148}$	4
A02+T3+E	468	$0.856_{0.172}$	$0.889_{0.167}$	$0.883_{0.175}$	2	$0.899_{0.161}$	$0.874_{0.188}$	3	$0.903_{0.146}$	$0.890_{0.160}$	1
A03+T1+E	1622	$0.778_{0.168}$	$0.845_{0.117}$	$0.827_{0.137}$	1	$0.854_{0.111}$	$0.824_{0.145}$	2	$0.881_{0.095}$	$0.823_{0.132}$	4
A03+T3+E	260	$0.891_{0.116}$	$0.912_{0.086}$	$0.876_{0.173}$	2	$0.923_{0.089}$	$0.868_{0.150}$	1	$0.932_{0.074}$	$0.874_{0.163}$	3
A04+T1+E	992	$0.850_{0.158}$	$0.880_{0.131}$	$0.860_{0.149}$	1	$0.888_{0.132}$	$0.866_{0.153}$	2	$0.906_{0.108}$	$0.856_{0.157}$	4
A04+T1+N	61	$0.760_{0.242}$	$0.840_{0.152}$	$0.823_{0.164}$	1	$0.837_{0.162}$	$0.786_{0.201}$	1	$0.827_{0.206}$	$0.789_{0.226}$	4
A04+T3+E	913	$0.912_{0.088}$	$0.939_{0.054}$	$0.934_{0.065}$	2	$0.948_{0.047}$	$0.926_{0.069}$	1	$0.951_{0.045}$	$0.932_{0.063}$	3
A04+T3+N	90	$0.877_{0.096}$	$0.910_{0.068}$	$0.905_{0.070}$	2	$0.928_{0.031}$	$0.908_{0.044}$	3	$0.926_{0.052}$	$0.913_{0.055}$	1
A05+T1+E	752	$0.815_{0.203}$	$0.862_{0.163}$	$0.837_{0.179}$	1	$0.873_{0.162}$	$0.827_{0.184}$	1	$0.882_{0.147}$	$0.841_{0.177}$	4
A05+T3+E	742	$0.875_{0.129}$	$0.903_{0.109}$	$0.891_{0.113}$	2	$0.916_{0.098}$	$0.878_{0.120}$	1	$0.919_{0.091}$	$0.891_{0.108}$	1
A06+T1+E	10	$0.824_{0.187}$	$0.902_{0.037}$	$0.885_{0.070}$	1	$0.909_{0.034}$	$0.889_{0.049}$	2	$0.909_{0.039}$	$0.880_{0.063}$	4
A06+T3+E	24	$0.862_{0.079}$	$0.916_{0.053}$	$0.916_{0.053}$	2	$0.934_{0.031}$	$0.923_{0.031}$	3	$0.933_{0.041}$	$0.929_{0.040}$	1
A07+T1+E	67	$0.820_{0.157}$	$0.877_{0.124}$	$0.867_{0.150}$	1	$0.890_{0.108}$	$0.862_{0.157}$	2	$0.897_{0.104}$	$0.862_{0.149}$	4
A07+T1+N	251	$0.837_{0.141}$	$0.892_{0.085}$	$0.879_{0.104}$	1	$0.903_{0.067}$	$0.875_{0.114}$	2	$0.905_{0.070}$	$0.873_{0.101}$	4
A07+T3+E	12	$0.925_{0.055}$	$0.938_{0.019}$	$0.937_{0.019}$	2	$0.939_{0.020}$	$0.916_{0.055}$	1	$0.947_{0.016}$	$0.932_{0.017}$	1
A07+T3+N	39	$0.863_{0.177}$	$0.918_{0.061}$	$0.913_{0.071}$	2	$0.933_{0.037}$	$0.899_{0.148}$	3	$0.934_{0.039}$	$0.914_{0.079}$	1
A08+T1+E	26	$0.666_{0.225}$	$0.750_{0.161}$	$0.680_{0.242}$	2	$0.747_{0.197}$	$0.653_{0.260}$	1	$0.793_{0.134}$	$0.666_{0.261}$	1
A08+T3+E	111	$0.605_{0.230}$	$0.668_{0.197}$	$0.626_{0.210}$	1	$0.677_{0.206}$	$0.628_{0.218}$	2	$0.735_{0.166}$	$0.669_{0.203}$	2
A09+T1+E	30	$0.815_{0.121}$	$0.841_{0.098}$	$0.784_{0.156}$	1	$0.873_{0.089}$	$0.833_{0.113}$	2	$0.884_{0.076}$	$0.812_{0.119}$	4
A09+T1+N	1	$0.953_{0.000}$	$0.927_{0.000}$	$0.927_{0.000}$	2	$0.955_{0.000}$	$0.955_{0.000}$	1	$0.947_{0.000}$	$0.947_{0.000}$	3
A09+T3+E	10	$0.900_{0.074}$	$0.918_{0.054}$	$0.918_{0.054}$	2	$0.933_{0.038}$	$0.909_{0.044}$	1	$0.937_{0.043}$	$0.919_{0.040}$	3
A09+T3+N	3	$0.894_{0.070}$	$0.911_{0.058}$	$0.911_{0.058}$	2	$0.957_{0.015}$	$0.957_{0.015}$	3	$0.944_{0.030}$	$0.944_{0.030}$	1

1. **Improved diversity without compromising quality**: for all *M* ≥ 2, choosing a single style that, for each annotator preference, maximizes agreement with the "ground truth" still outperforms 1-StyleSeg.

2. **Performance improves as *M* increases.**

3. **Ability to learn tool-specific latent factors:** Without specifically training for it, a 3-StyleSeg model is able to choose a unique style for each of the three tools ("T1", "T2", "T3").

Quantifying Annotator-Style Alignment: A New Measure

If we model 3 styles, the best style can be the one that

- best matches 100% of images (**perfect alignment**), or
- best matches, say, 34% of images (**weak alignment**).

Quantifying Annotator-Style Alignment: A New Measure

If we model 3 styles, the best style can be the one that

- best matches 100% of images (**perfect alignment**), or
- best matches, say, 34% of images (**weak alignment**).

How do we quantify this annotator-style alignment strength?

Quantifying Annotator-Style Alignment: A New Measure

If we model 3 styles, the best style can be the one that

- best matches 100% of images (**perfect alignment**), or
- best matches, say, 34% of images (**weak alignment**).

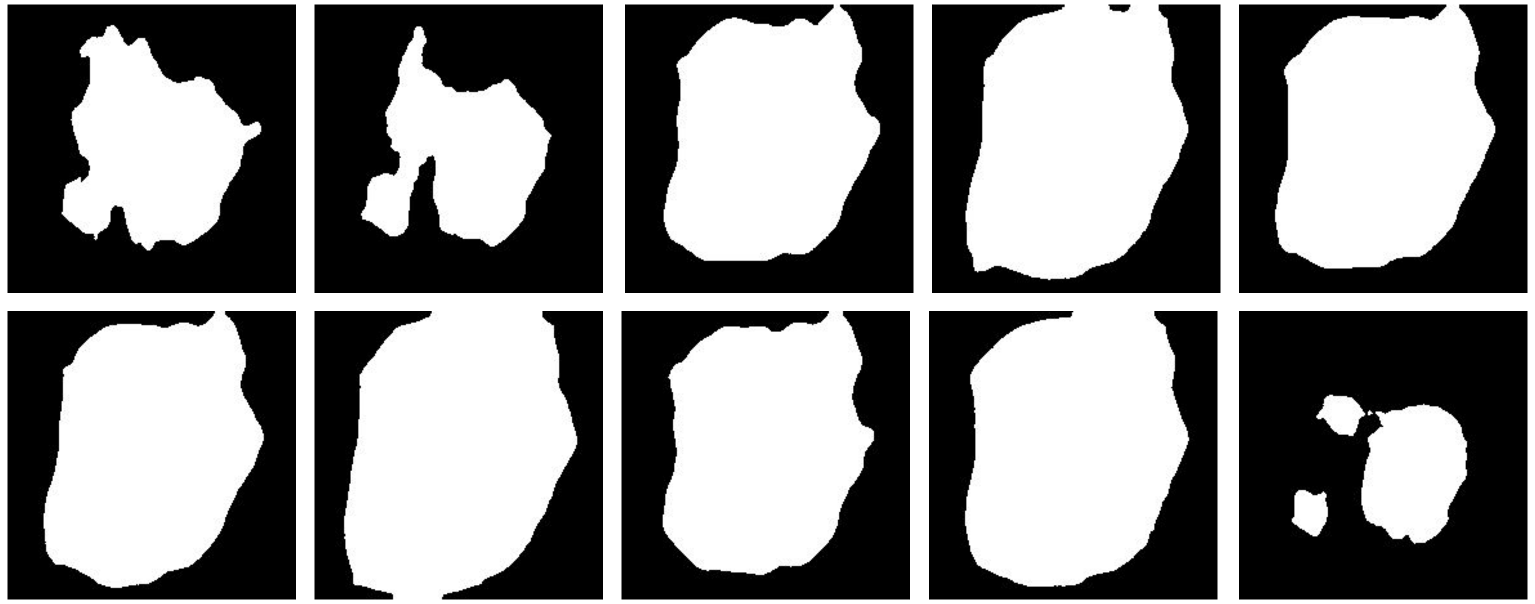
How do we quantify this annotator-style alignment strength?

$$\mathrm{AS}^2 = 1 - \frac{-\sum_{i=1}^{M} q_i \log_2 q_i}{-\sum_{j=1}^{M} \frac{1}{M} \log_2 \frac{1}{M}}$$

Quantifying Annotator-Style Alignment: A New Measure

If we model 3 styles, the best style can be the one that

- best matches 100% of images (**perfect alignment**), or
- best matches, say, 34% of images (**weak alignment**).



How do we quantify this annotator-style alignment strength?

$$\mathrm{AS}^2 = 1 - \frac{-\sum_{i=1}^{M} q_i \log_2 q_i}{-\sum_{j=1}^{M} \frac{1}{M} \log_2 \frac{1}{M}}$$

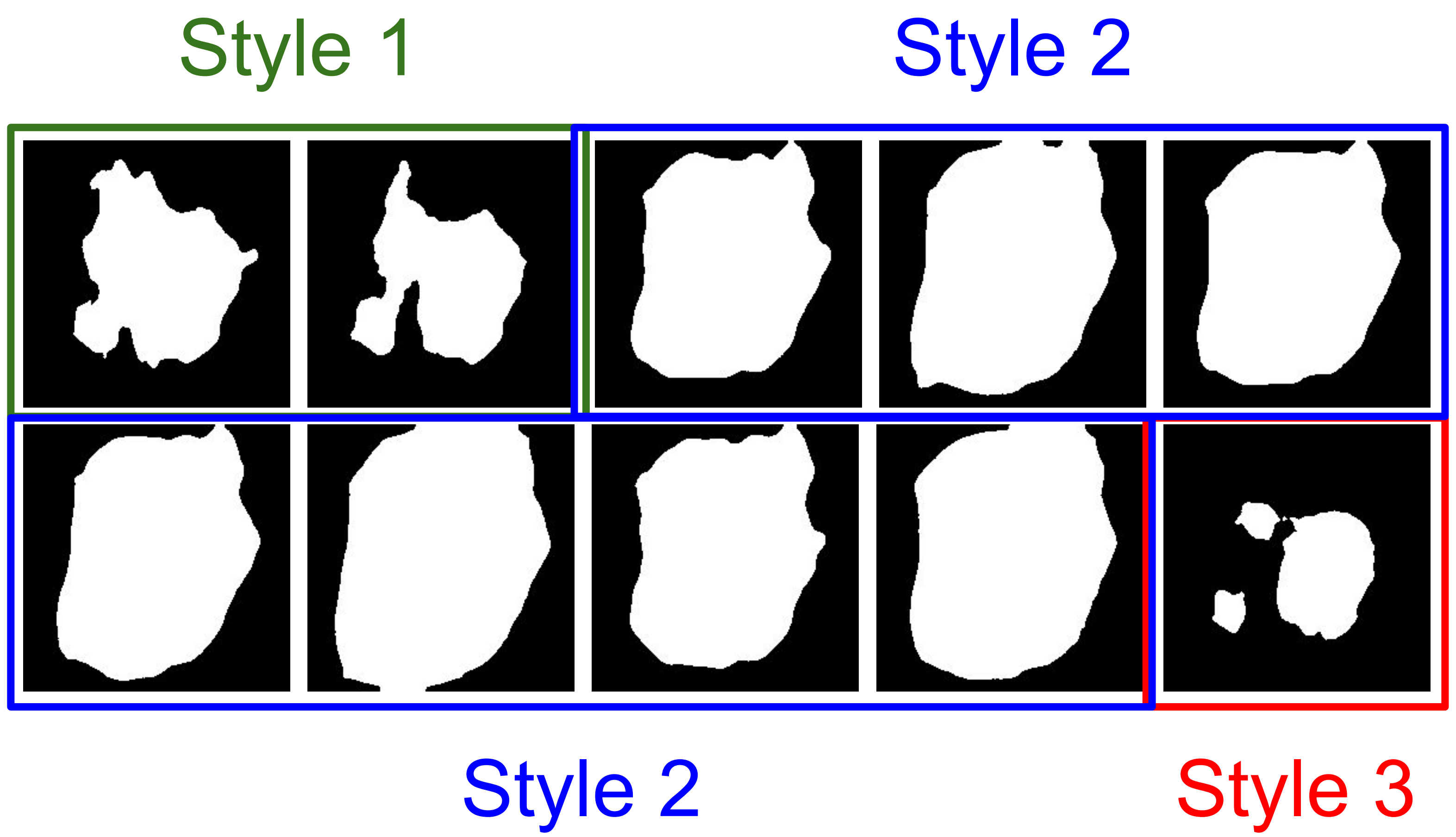
Quantifying Annotator-Style Alignment: A New Measure

If we model 3 styles, the best style can be the one that

- best matches 100% of images (**perfect alignment**), or
- best matches, say, 34% of images (**weak alignment**).

How do we quantify this annotator-style alignment strength?

$$\mathrm{AS}^2 = 1 - \frac{-\sum_{i=1}^{M} q_i \log_2 q_i}{-\sum_{j=1}^{M} \frac{1}{M} \log_2 \frac{1}{M}}$$



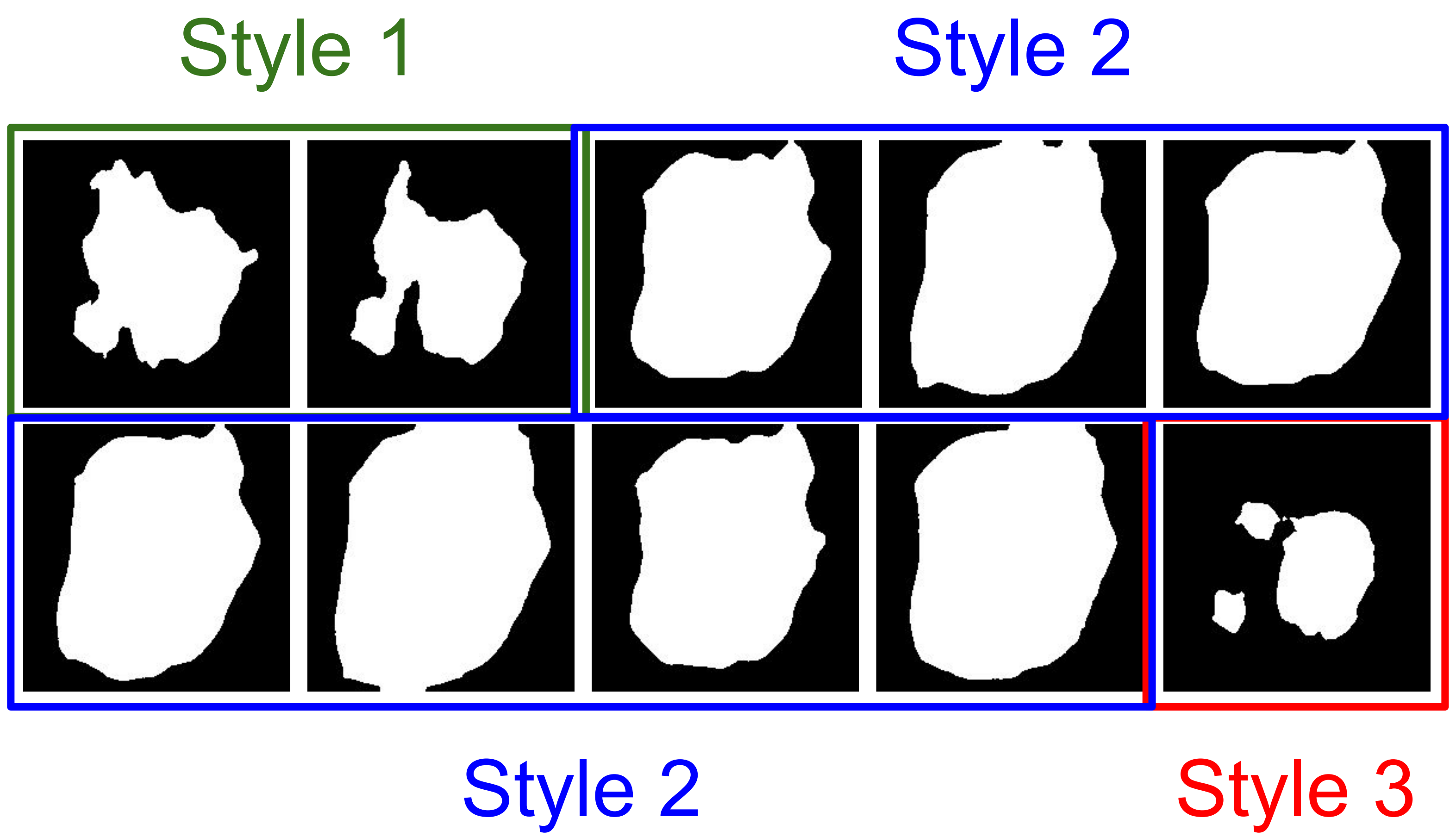
Quantifying Annotator-Style Alignment: A New Measure

If we model 3 styles, the best style can be the one that

- best matches 100% of images (**perfect alignment**), or
- best matches, say, 34% of images (**weak alignment**).

How do we quantify this <u>annotator-style alignment strength</u>?

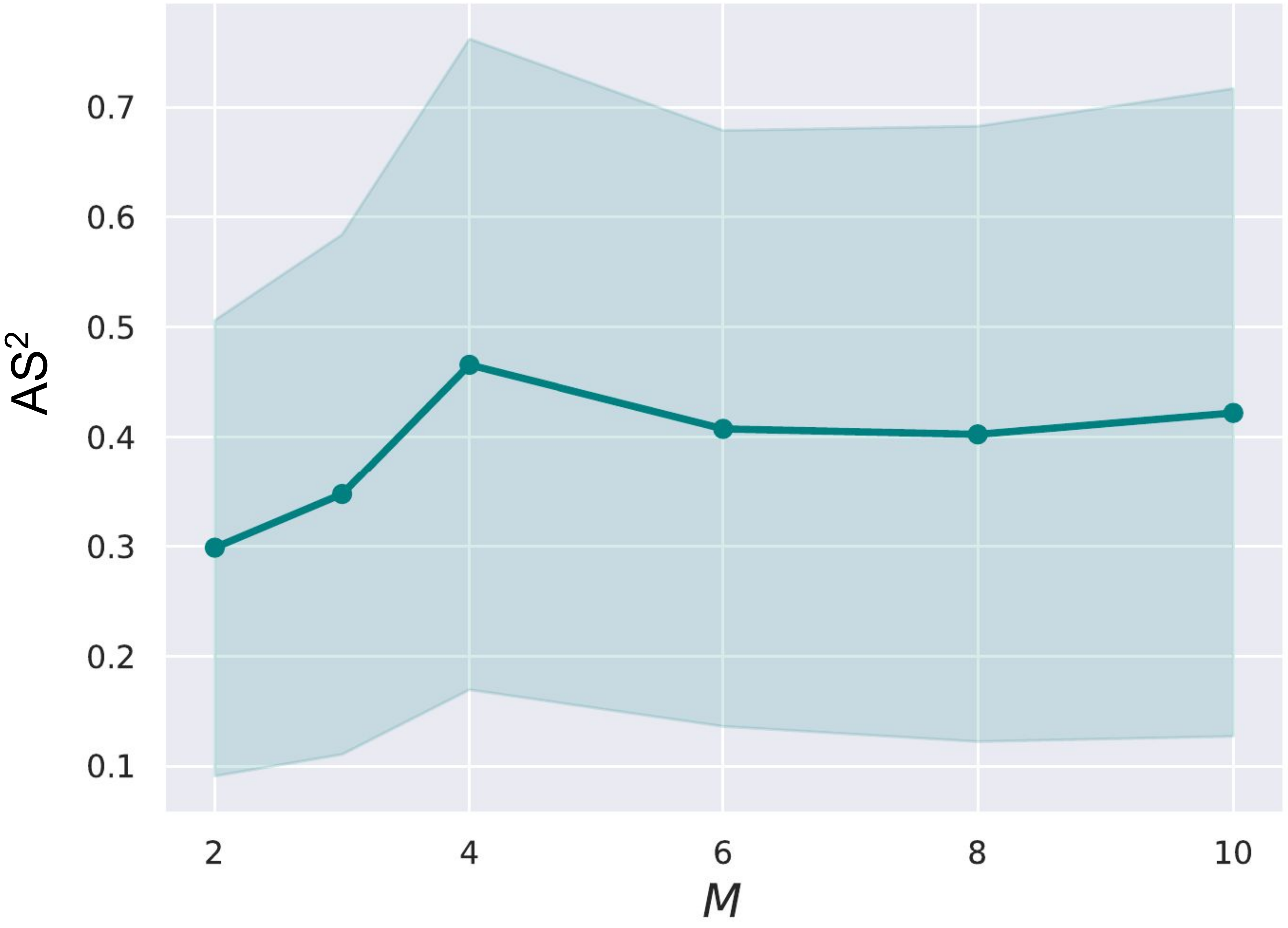
$$\mathrm{AS}^2 = 1 - \frac{-\sum_{i=1}^{M} q_i \log_2 q_i}{-\sum_{j=1}^{M} \frac{1}{M} \log_2 \frac{1}{M}}$$



$q_1 = 0.2$, $q_2 = 0.7$, $q_3 = 0.1$

$q = [0.2, 0.7, 0.1] \Rightarrow \mathrm{AS}^2 = 0.27.$

Quantifying Annotator-Style Alignment



Modeling **more styles** captures **more diversity** and is not detrimental to segmentation quality.

Conclusion

- Formulated the **problem of segmentation style discovery**, and showed that StyleSeg discovers styles that are **plausible**, **diverse**, and **semantically consistent**.

Conclusion

- Formulated the **problem of segmentation style discovery**, and showed that StyleSeg discovers styles that are **plausible**, **diverse**, and **semantically consistent**.
- The **largest multi-annotator SLS dataset** (> 13.5k image-mask pairs) with **annotator correspondence** curated from the ISIC Archive.

Conclusion

- Formulated the **problem of segmentation style discovery**, and showed that StyleSeg discovers styles that are **plausible**, **diverse**, and **semantically consistent**.
- The **largest multi-annotator SLS dataset** (> 13.5k image-mask pairs) with **annotator correspondence** curated from the ISIC Archive.
- **A new measure** for **quantifying the strength of alignment** between annotators' preferences and styles.

Conclusion

- Formulated the **problem of segmentation style discovery**, and showed that StyleSeg discovers styles that are **plausible**, **diverse**, and **semantically consistent**.

- The **largest multi-annotator SLS dataset** (> 13.5k image-mask pairs) with **annotator correspondence** curated from the ISIC Archive.

- **A new measure** for **quantifying the strength of alignment** between annotators' preferences and styles.

- **Future work** may look at approaches to **finding the optimal number of styles** in a segmentation dataset.

References

[1] Silletti et al., “Variability in human and automatic segmentation of melanocytic lesions”, *EMBC*, 2009.

[2] Mirikharaji et al., “D-LEMA: Deep learning ensembles from multiple annotations-application to skin lesion segmentation”, *CVPR ISIC* 2021.

[3] Ribeiro et al., “Less is more: Sample selection and label conditioning improve skin lesion segmentation”, *CVPR ISIC* 2020.

[4] Rupprecht et al., “Learning in an uncertain world: Representing ambiguity through multiple hypotheses”, *ICCV* 2017.

Thank you.

Questions?

kabhishe@sfu.ca

Acknowledgements

NSERC CRSNG

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

NVIDIA