Shape Index Descriptors Applied to Texture-Based Galaxy Analysis — Supplementary Material —

Kim Steenstrup Pedersen†, Kristoffer Stensbo-Smidt†, Andrew Zirm‡, Christian Igel†
† Department of Computer Science, and ‡ Dark Cosmology Centre, Niels Bohr Institute
University of Copenhagen, Denmark
http://image.diku.dk/MLLab/SkyML.php

Figure 1, figure 3, and table 1 refer to the analysis where only galaxies that pass the selection criteria of all masks are considered. These give additional results on top of the results reported in table 1 of the paper. Figure 2, figure 4, and table 2 refer to the analysis where all galaxies that pass the selection criteria of a given mask are considered. The data sets for all masks are larger than the data set used in table 1. Finally, in figure 5 we illustrate the steps for processing the segmentation mask prior to feature extraction.

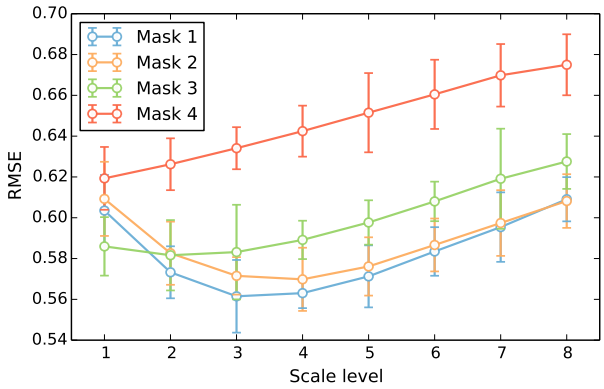


Figure 1: Plot of RMSE (error bars indicate 1 standard deviation of the CV error) of Linear *gri* (SI) across the 8 scale levels for the four masks. The plot is identical to Fig. 2 in the paper.

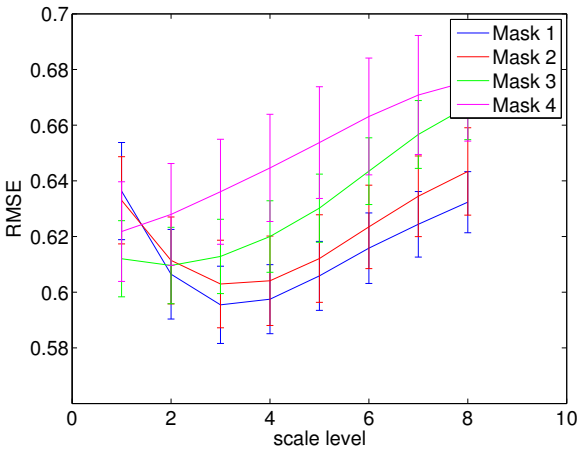
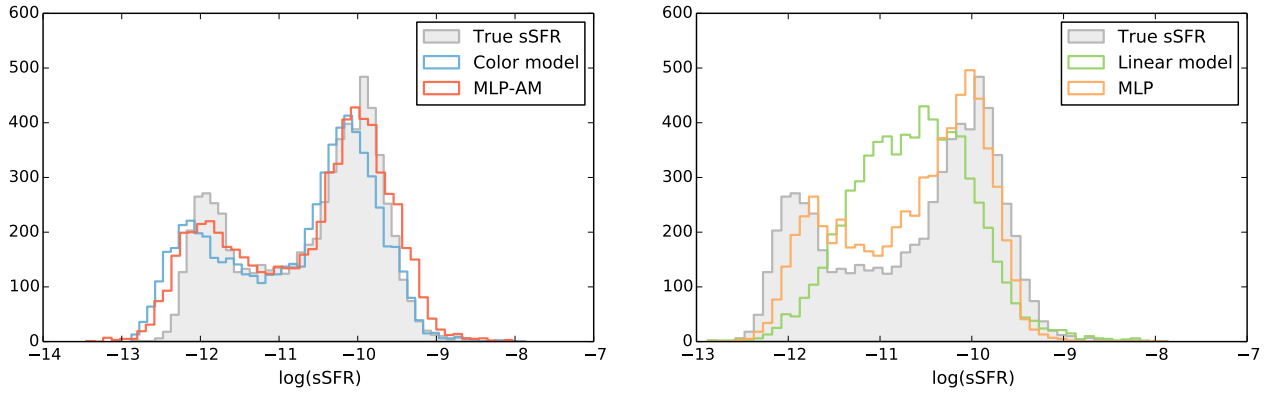
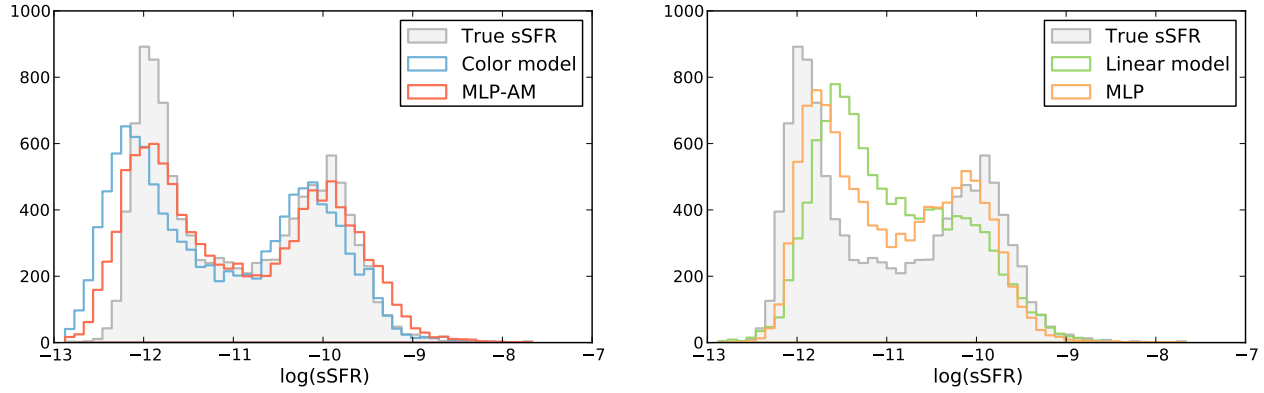


Figure 2: Plot of RMSE (error bars indicate 1 standard deviation of the CV error) of Linear *gri* (SI) across the 8 scale levels for the four masks. The plot is identical to Fig. 2 in the paper, but here all galaxies that pass the selection criteria of mask 1 are considered (see amount of galaxies in table 2).



(a) (b)

Figure 3: Plots of the distributions of predicted sSFR values for different predictors and the ground truth for mask 1, using the *gri* and shape index (SI) features. The plots show the same as Fig. 3 in the paper, but the plot has been split into two to make it easier to distinguish the distributions.



(a) (b)

Figure 4: Plots of the distributions of predicted sSFR values for different predictors and the ground truth for mask 1, using the *gri* and shape index (SI) features. The plots are identical to those in Fig. 3, but all galaxies that pass the selection criteria in mask 1 are used (see amount of galaxies in table 2).

Table 1: Extension of table 1 in the paper, summarising our results for different model-feature pairs applied to either single bands (g, r, and i) or all bands (gri) using four different masks. The results are based on 6880 images passing the inclusion criteria. The numbers in the table indicate RMSE and cross validation standard deviation. *Average* refers to predicting the training data mean and *Color* is the current state-of-the-art physical model. *Linear* and *MLP* denote linear and non-linear regression, and *Linear-AM* and *MLP-AM* are the additive model counterparts. Gradient orientation (*GO*) and shape index (*SI*) features as well as their combination (*All*) and second order features (*2nd*) are considered.

Method	**Band (features)**	**Mask 1**	**Mask 2**	**Mask 3**	**Mask 4**
Average		0.88 ± 0.02	0.88 ± 0.01	0.88 ± 0.01	0.88 ± 0.01
Color		0.33 ± 0.01	0.33 ± 0.02	0.33 ± 0.02	0.33 ± 0.02
Linear	*g* (all)	0.61 ± 0.01	0.62 ± 0.02	0.62 ± 0.01	0.65 ± 0.01
	r (all)	0.65 ± 0.02	0.63 ± 0.02	0.63 ± 0.01	0.67 ± 0.02
	i (all)	0.65 ± 0.02	0.64 ± 0.02	0.64 ± 0.02	0.67 ± 0.01
	gri (all)	0.53 ± 0.02	0.54 ± 0.02	0.55 ± 0.02	0.59 ± 0.02
Linear	*g* (SI)	0.62 ± 0.01	0.63 ± 0.01	0.62 ± 0.01	0.65 ± 0.01
	r (SI)	0.66 ± 0.02	0.64 ± 0.01	0.64 ± 0.02	0.67 ± 0.02
	i (SI)	0.66 ± 0.02	0.64 ± 0.02	0.64 ± 0.01	0.67 ± 0.01
	gri (SI)	0.53 ± 0.02	0.54 ± 0.02	0.55 ± 0.02	0.59 ± 0.01
Linear	*g* (GO)	0.86 ± 0.02	0.87 ± 0.01	0.86 ± 0.01	0.85 ± 0.02
	r (GO)	0.87 ± 0.02	0.86 ± 0.01	0.85 ± 0.01	0.85 ± 0.01
	i (GO)	0.87 ± 0.02	0.86 ± 0.01	0.85 ± 0.01	0.85 ± 0.01
	gri (GO)	0.81 ± 0.02	0.83 ± 0.01	0.84 ± 0.01	0.85 ± 0.02
MLP	*g* (all)	0.55 ± 0.01	0.57 ± 0.02	0.58 ± 0.02	0.61 ± 0.01
	r (all)	0.61 ± 0.02	0.59 ± 0.02	0.61 ± 0.02	0.63 ± 0.01
	i (all)	0.61 ± 0.02	0.60 ± 0.02	0.61 ± 0.01	0.64 ± 0.02
	gri (all)	0.49 ± 0.02	0.50 ± 0.01	0.52 ± 0.01	0.55 ± 0.02
MLP	*g* (SI)	0.55 ± 0.02	0.58 ± 0.01	0.59 ± 0.01	0.61 ± 0.02
	r (SI)	0.61 ± 0.02	0.60 ± 0.02	0.61 ± 0.01	0.63 ± 0.03
	i (SI)	0.61 ± 0.02	0.61 ± 0.01	0.62 ± 0.02	0.63 ± 0.03
	gri (SI)	0.50 ± 0.02	0.50 ± 0.01	0.52 ± 0.01	0.56 ± 0.01
MLP	*g* (GO)	0.83 ± 0.02	0.84 ± 0.02	0.83 ± 0.02	0.82 ± 0.02
	r (GO)	0.85 ± 0.01	0.84 ± 0.02	0.83 ± 0.02	0.82 ± 0.02
	i (GO)	0.85 ± 0.02	0.84 ± 0.02	0.83 ± 0.01	0.82 ± 0.02
	gri (GO)	0.78 ± 0.02	0.79 ± 0.02	0.80 ± 0.01	0.80 ± 0.02
Linear-AM	*g* (all)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
	r (all)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
	i (all)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
	gri (all)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
Linear-AM	*g* (SI)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
	r (SI)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
	i (SI)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
	gri (SI)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
MLP-AM	*g* (all)	0.29 ± 0.02	0.29 ± 0.03	0.29 ± 0.02	0.29 ± 0.02
	r (all)	0.29 ± 0.02	0.29 ± 0.03	0.29 ± 0.02	0.29 ± 0.02
	i (all)	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.02
	gri (all)	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02

Continued on next page...

Method	Band (features)		Mask 1	Mask 2	Mask 3	Mask 4
MLP-AM	*g*	(SI)	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02
	r	(SI)	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.03
	i	(SI)	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.03
	gri	(SI)	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.02
Linear	*g*	(2nd)	0.69 ± 0.06	0.75 ± 0.10	0.83 ± 0.38	0.77 ± 0.27
	r	(2nd)	0.71 ± 0.07	0.73 ± 0.07	0.74 ± 0.08	0.80 ± 0.21
	i	(2nd)	0.72 ± 0.07	0.68 ± 0.04	0.73 ± 0.11	0.72 ± 0.03
	gri	(2nd)	0.53 ± 0.03	0.57 ± 0.05	0.68 ± 0.31	0.64 ± 0.05
MLP	*g*	(2nd)	0.63 ± 0.02	0.69 ± 0.05	27.20 ± 78.54	0.73 ± 0.07
	r	(2nd)	0.68 ± 0.04	0.75 ± 0.22	0.78 ± 0.17	48.18 ± 142.37
	i	(2nd)	0.68 ± 0.03	0.67 ± 0.03	1.15 ± 1.14	0.81 ± 0.23
	gri	(2nd)	0.54 ± 0.06	0.54 ± 0.02	0.67 ± 0.22	28.54 ± 83.81

Table 2: A summary of our results for different model-feature pairs applied to either single bands (g, r, and i) or all bands (gri) using four different masks. The total number of galaxy images included in the analysis is indicated for each mask. The numbers in the table indicate RMSE and cross validation standard deviation. *Average* refers to predicting the training data mean and *Color* is the current state-of-the-art physical model. *Linear* and *MLP* denote linear and non-linear regression, and *Linear-AM* and *MLP-AM* are the additive model counterparts. Gradient orientation (*GO*) and shape index (*SI*) features as well as their combination *all* are considered.

Method	**Band (features)**		**Mask 1** 11 450 galaxies	**Mask 2** 11 198 galaxies	**Mask 3** 10 492 galaxies	**Mask 4** 7291 galaxies
Average			0.93 ± 0.01	0.93 ± 0.01	0.93 ± 0.01	0.89 ± 0.02
Color			0.36 ± 0.01	0.36 ± 0.01	0.35 ± 0.01	0.34 ± 0.01
Linear	*r*	(all)	0.67 ± 0.02	0.65 ± 0.01	0.64 ± 0.01	0.67 ± 0.02
	g	(all)	0.63 ± 0.01	0.64 ± 0.01	0.63 ± 0.01	0.65 ± 0.01
	i	(all)	0.68 ± 0.01	0.66 ± 0.01	0.65 ± 0.01	0.67 ± 0.01
	gri	(all)	0.56 ± 0.01	0.56 ± 0.01	0.57 ± 0.01	0.59 ± 0.02
Linear	*r*	(SI)	0.68 ± 0.02	0.66 ± 0.02	0.64 ± 0.01	0.67 ± 0.02
	g	(SI)	0.64 ± 0.01	0.64 ± 0.01	0.63 ± 0.01	0.65 ± 0.01
	i	(SI)	0.68 ± 0.01	0.66 ± 0.02	0.65 ± 0.01	0.67 ± 0.01
	gri	(SI)	0.56 ± 0.02	0.57 ± 0.01	0.57 ± 0.01	0.59 ± 0.02
Linear	*r*	(GO)	0.91 ± 0.01	0.88 ± 0.02	0.87 ± 0.01	0.85 ± 0.01
	g	(GO)	0.90 ± 0.01	0.89 ± 0.01	0.87 ± 0.02	0.85 ± 0.01
	i	(GO)	0.90 ± 0.01	0.88 ± 0.01	0.87 ± 0.01	0.85 ± 0.01
	gri	(GO)	0.86 ± 0.01	0.87 ± 0.02	0.86 ± 0.01	0.85 ± 0.02
MLP	*r*	(all)	0.63 ± 0.02	0.62 ± 0.02	0.62 ± 0.02	0.63 ± 0.02
	g	(all)	0.58 ± 0.01	0.60 ± 0.01	0.60 ± 0.01	0.61 ± 0.02
	i	(all)	0.64 ± 0.01	0.62 ± 0.02	0.62 ± 0.01	0.63 ± 0.02
	gri	(all)	0.52 ± 0.02	0.54 ± 0.01	0.54 ± 0.01	0.56 ± 0.02
MLP	*r*	(SI)	0.63 ± 0.02	0.63 ± 0.02	0.62 ± 0.02	0.63 ± 0.01
	g	(SI)	0.58 ± 0.02	0.61 ± 0.01	0.60 ± 0.01	0.61 ± 0.02
	i	(SI)	0.64 ± 0.01	0.63 ± 0.02	0.63 ± 0.02	0.63 ± 0.01
	gri	(SI)	0.52 ± 0.02	0.54 ± 0.01	0.54 ± 0.02	0.56 ± 0.01
MLP	*r*	(GO)	0.87 ± 0.01	0.85 ± 0.01	0.83 ± 0.02	0.81 ± 0.03
	g	(GO)	0.86 ± 0.01	0.86 ± 0.01	0.83 ± 0.01	0.82 ± 0.02
	i	(GO)	0.87 ± 0.02	0.86 ± 0.01	0.82 ± 0.02	0.81 ± 0.02
	gri	(GO)	0.81 ± 0.02	0.82 ± 0.02	0.80 ± 0.02	0.80 ± 0.02
Linear-AM	*r*	(all)	0.30 ± 0.01	0.30 ± 0.01	0.30 ± 0.01	0.29 ± 0.01
	g	(all)	0.30 ± 0.01	0.30 ± 0.01	0.30 ± 0.02	0.29 ± 0.01
	i	(all)	0.30 ± 0.02	0.29 ± 0.01	0.29 ± 0.01	0.29 ± 0.02
	gri	(all)	0.30 ± 0.01	0.30 ± 0.01	0.30 ± 0.01	0.30 ± 0.01
Linear-AM	*r*	(SI)	0.30 ± 0.01	0.30 ± 0.01	0.30 ± 0.01	0.29 ± 0.01
	g	(SI)	0.30 ± 0.01	0.30 ± 0.01	0.30 ± 0.02	0.29 ± 0.01
	i	(SI)	0.30 ± 0.02	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.01
	gri	(SI)	0.30 ± 0.01	0.29 ± 0.01	0.30 ± 0.01	0.29 ± 0.01
MLP-AM	*r*	(all)	0.30 ± 0.01	0.29 ± 0.02	0.30 ± 0.01	0.29 ± 0.02
	g	(all)	0.30 ± 0.01	0.30 ± 0.01	0.29 ± 0.01	0.29 ± 0.02
	i	(all)	0.30 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
	gri	(all)	0.29 ± 0.01	0.29 ± 0.01	0.29 ± 0.01	0.29 ± 0.02
MLP-AM	*r*	(SI)	0.30 ± 0.01	0.30 ± 0.01	0.29 ± 0.01	0.29 ± 0.01
	g	(SI)	0.30 ± 0.01	0.30 ± 0.02	0.29 ± 0.01	0.29 ± 0.02
	i	(SI)	0.30 ± 0.02	0.29 ± 0.01	0.29 ± 0.01	0.29 ± 0.02
	gri	(SI)	0.29 ± 0.01	0.29 ± 0.01	0.29 ± 0.01	0.29 ± 0.01

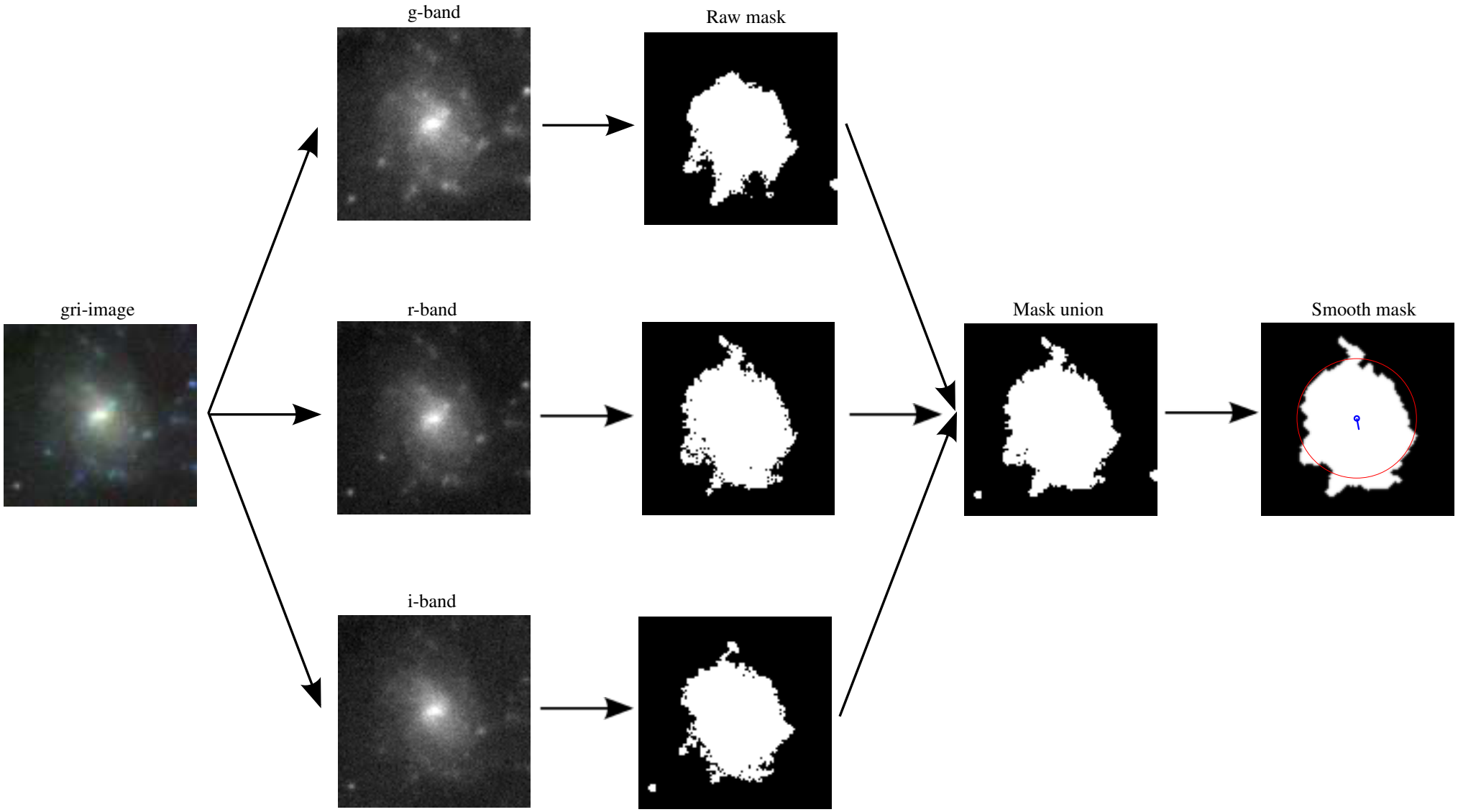


Figure 5: Illustration of the preprocessing steps on the segmentation masks: (Left) The original *gri*-image and the individual band images. (Middle) the raw masks (mask 1) for each of the bands and the union of these masks. (Right) Final smoothed mask and the estimated fiducial orientation (blue line) and the Petrosian radius visualized as a red circle centered in the center of mass of the galaxy.