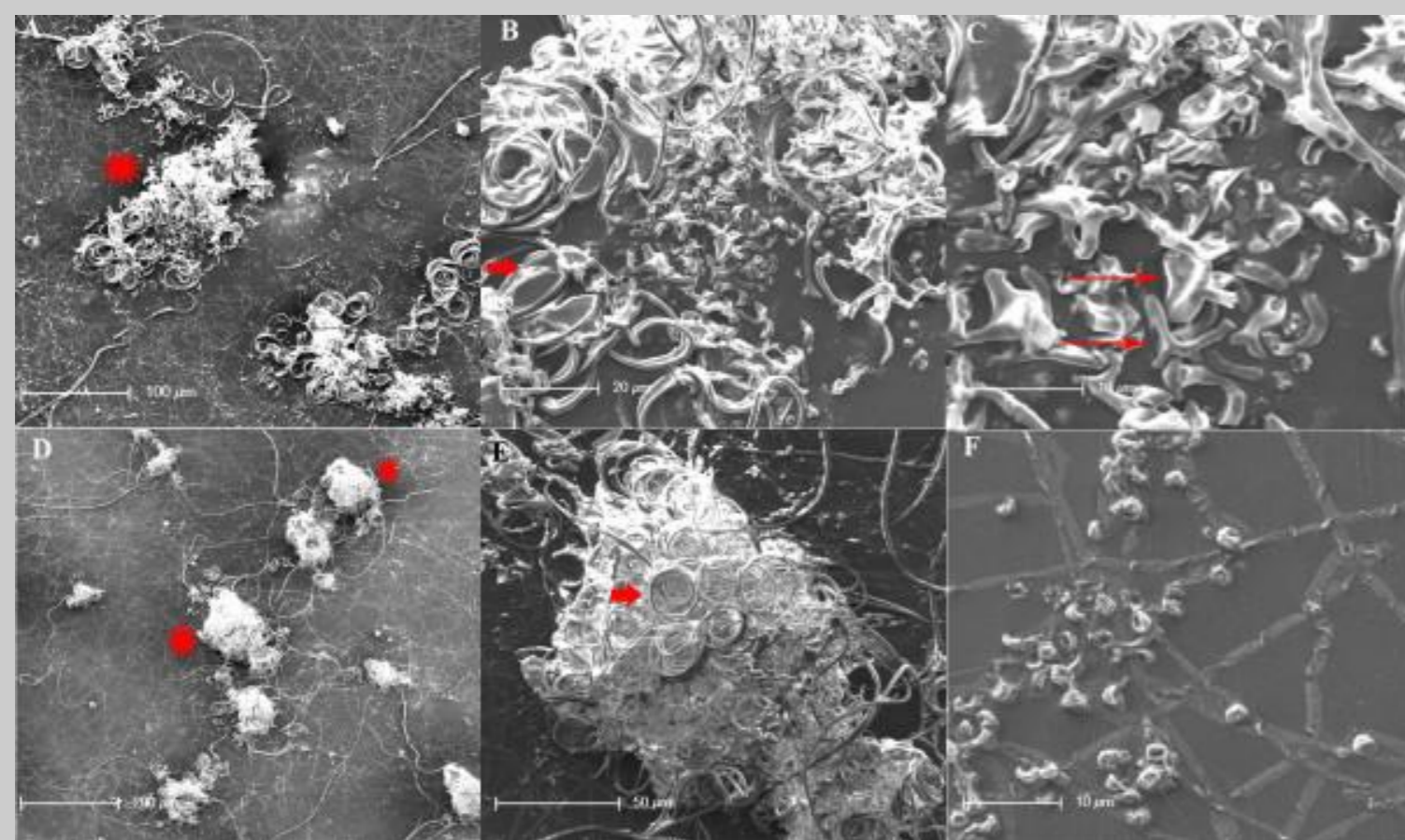
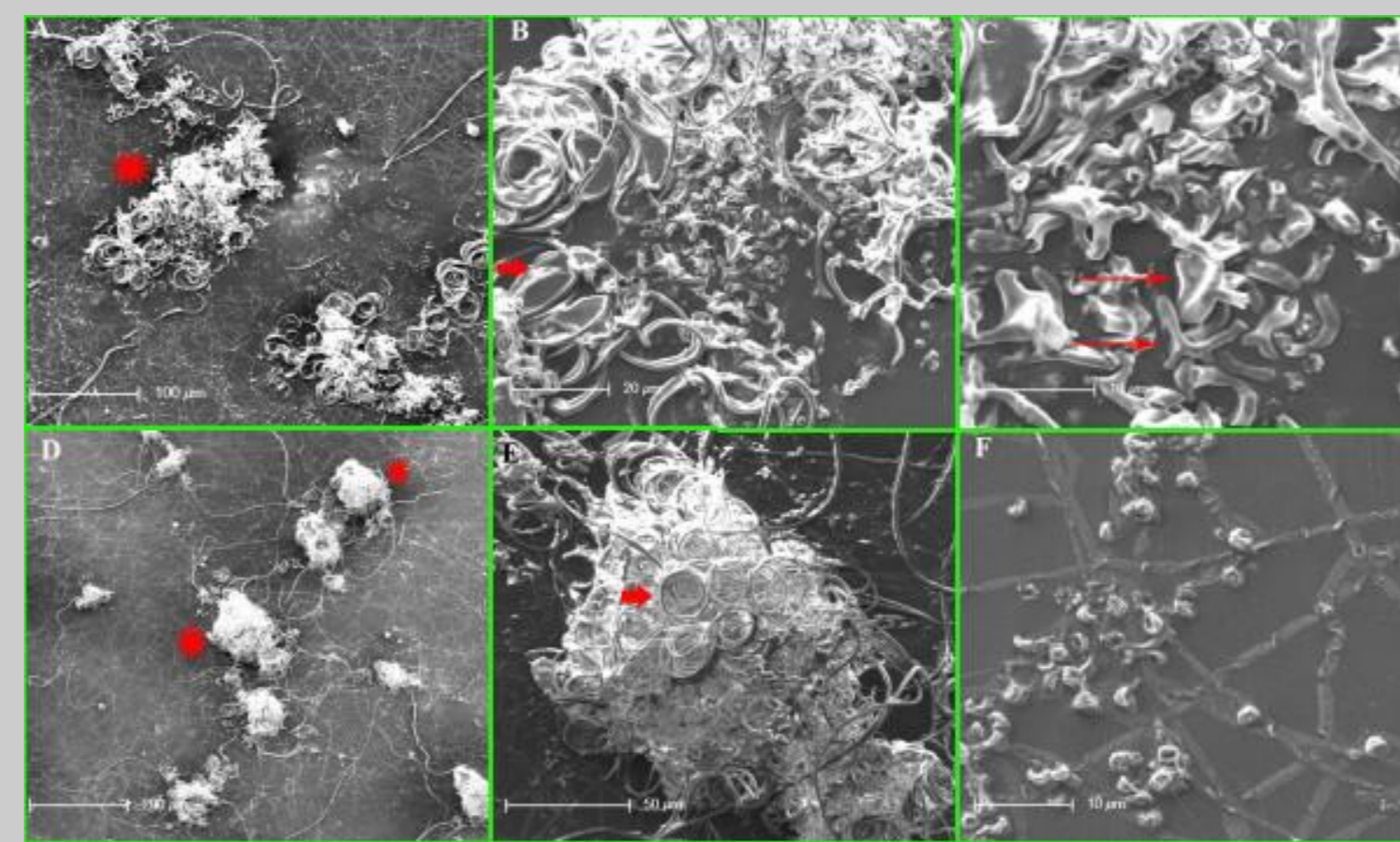




Problem statement



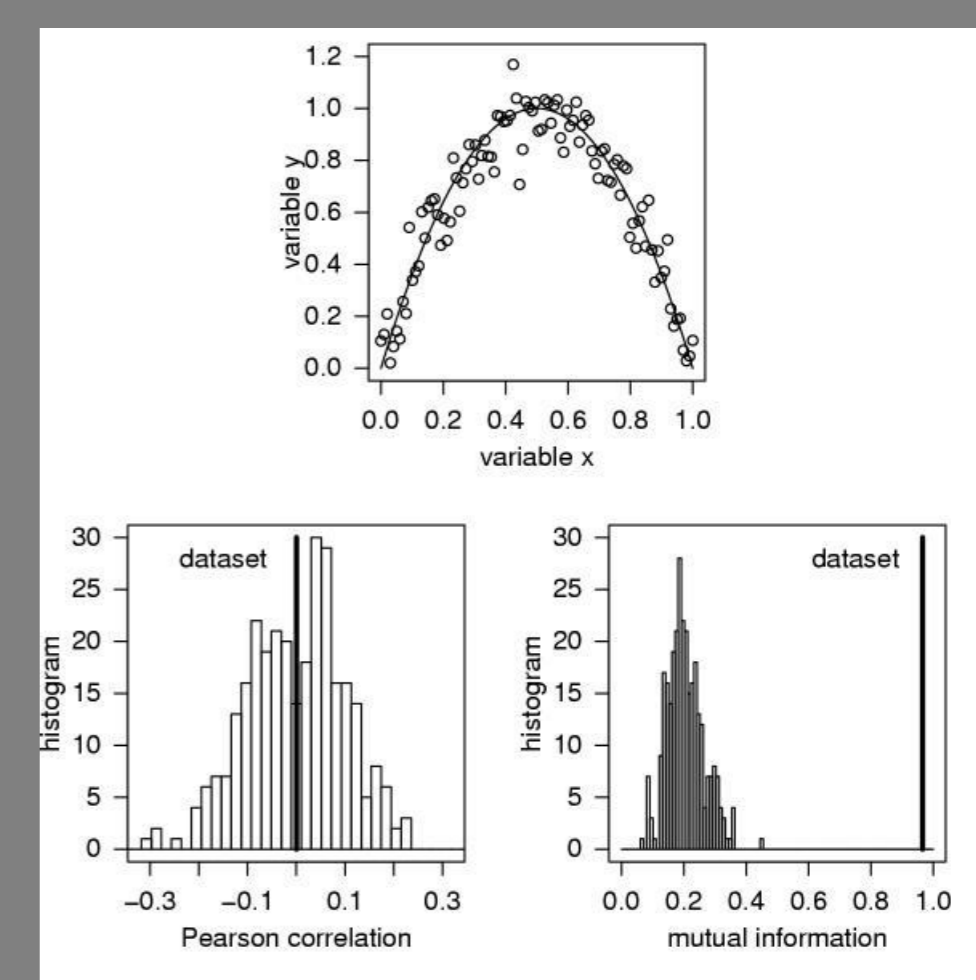
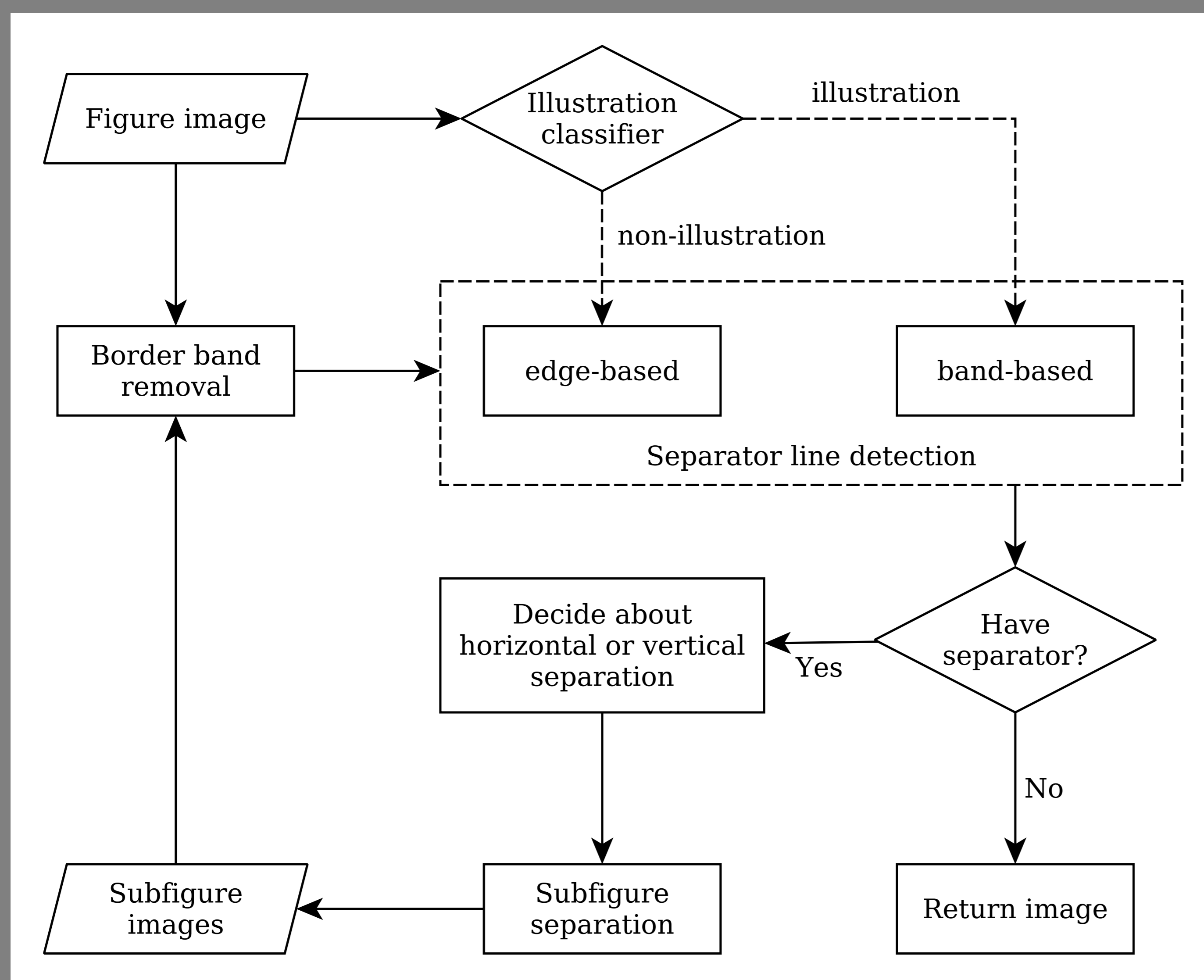
automatic
separation



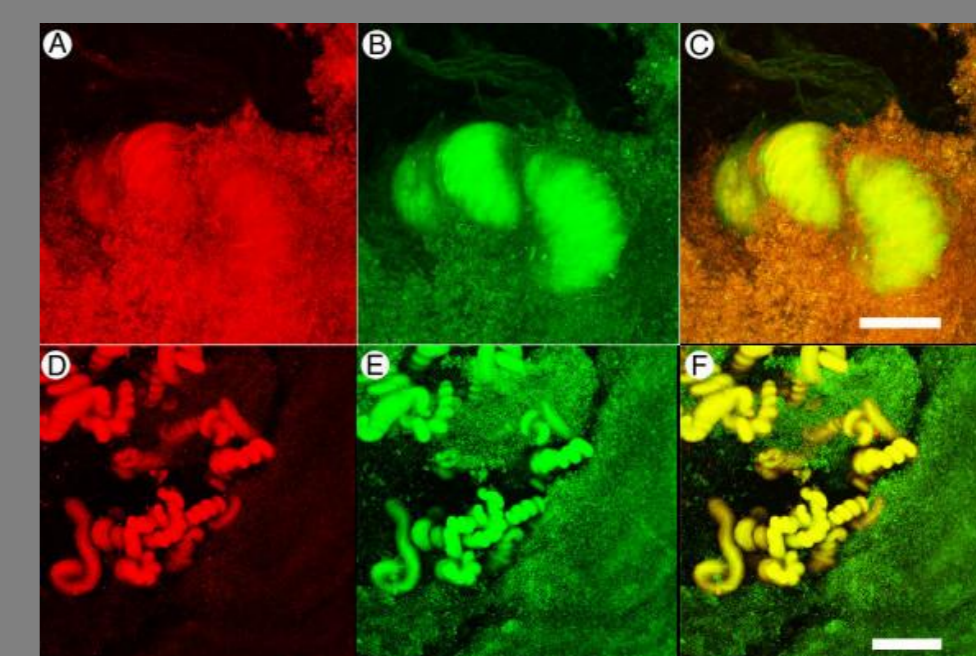
Compound image

segmented into subfigures

Proposed algorithm



Illustration



Non-illustration

Evaluation

Dataset	images	illustrations
Illustration classifier training	1071	57%
Compound figure separation	3381	54%

Compound figure separation	Accuracy	
	submitted	optimized*
band-based	30%	69%
edge-based	35%	55%
proposed	49%	73%
best competitor (NLM)	85%	

* Parameter optimization using the training set was done only after paper submission.

Further work

- Improve training set quality for classifier
- Optimize decisions of illustration classifier
- Eliminate small bounding boxes
- Remove image markup
- Extract image labels