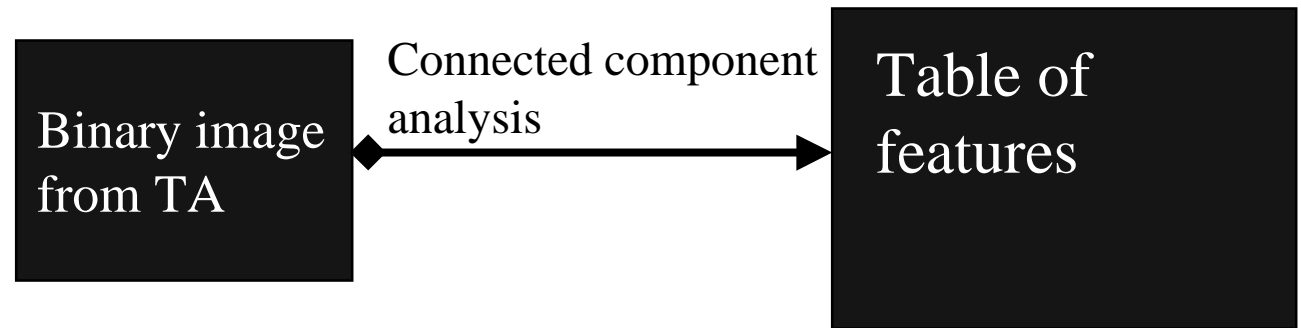




Your project submission



Training your classifier

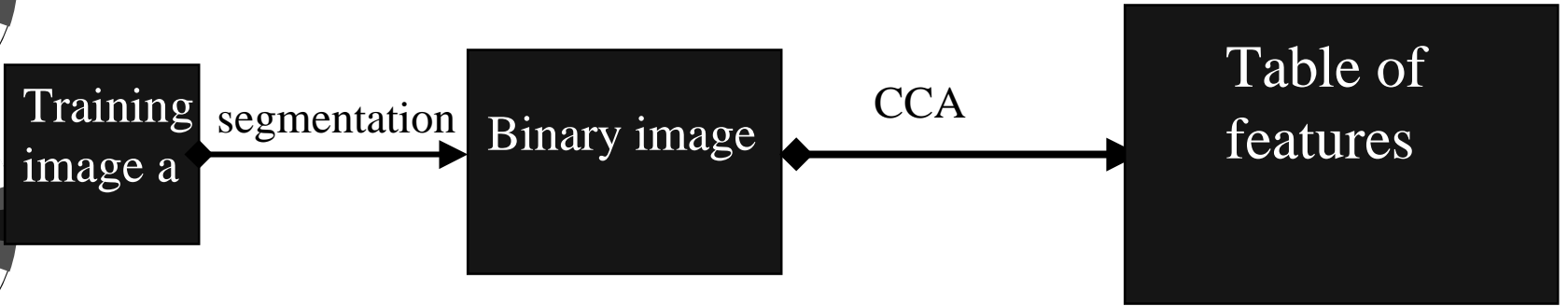
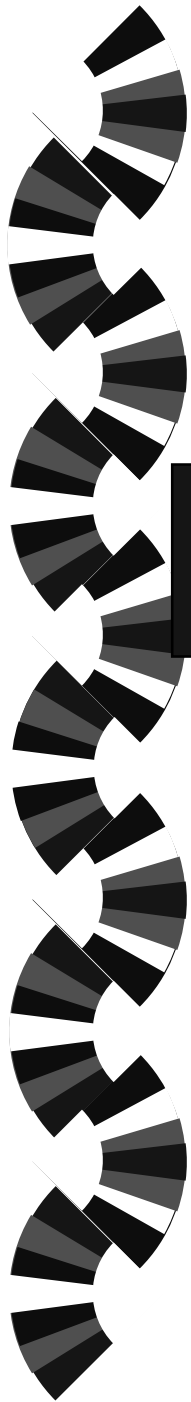
Training
image a

segmentation

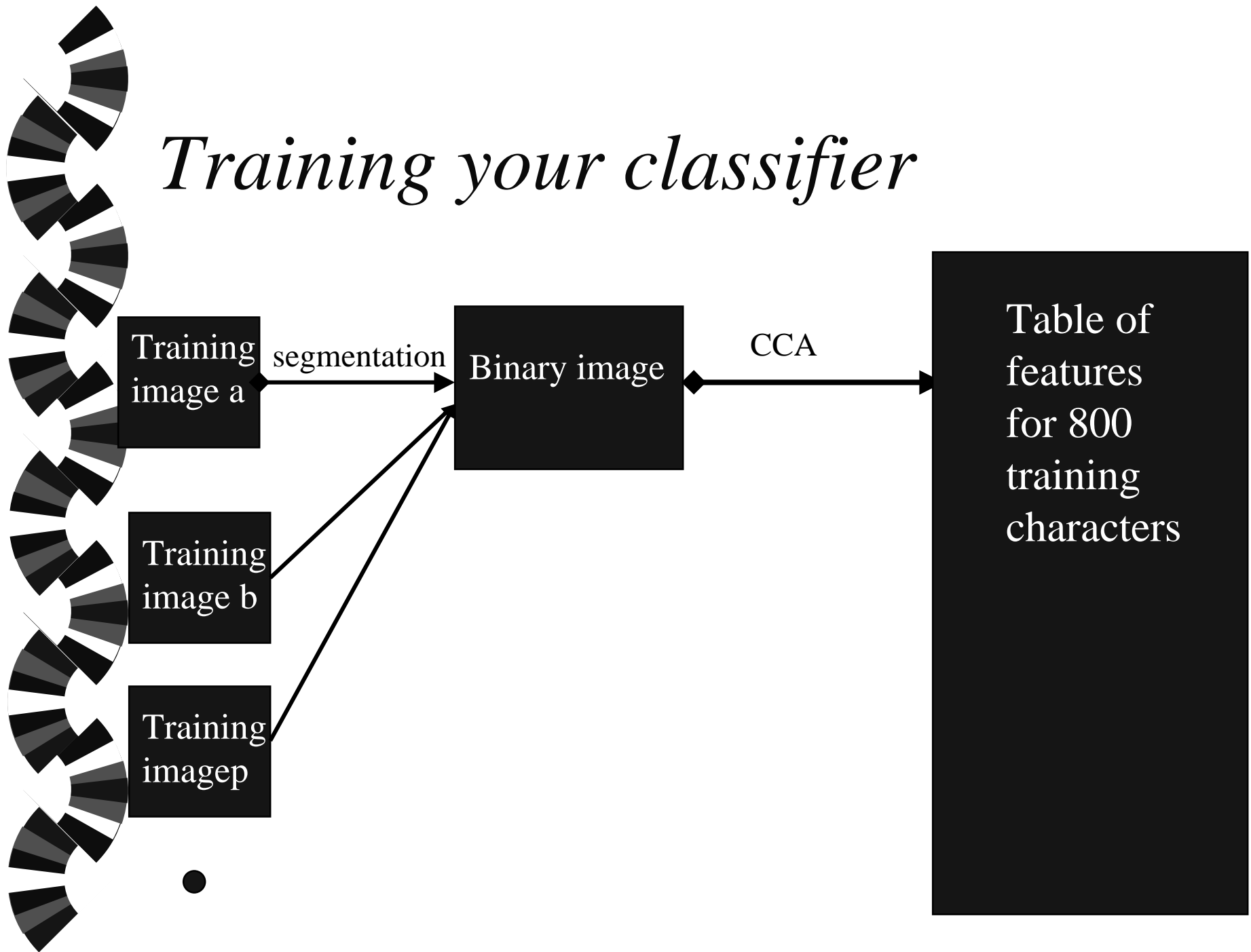
Binary image

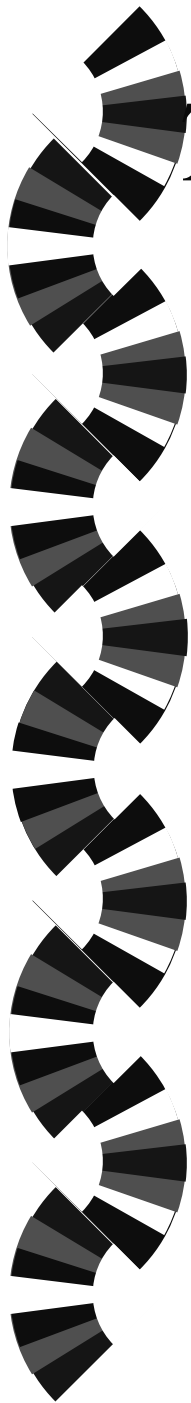
CCA

Table of
features



Training your classifier





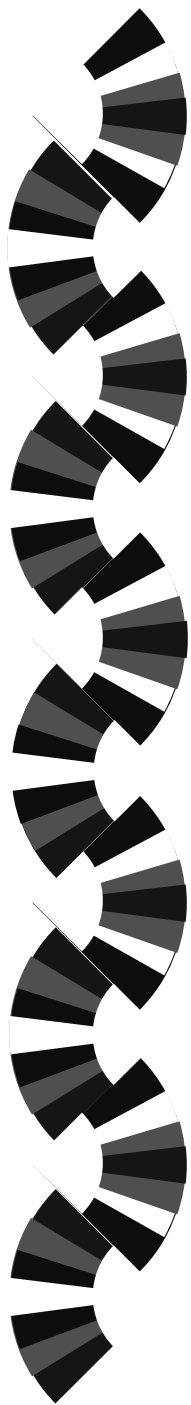
Training the classifier

- For each subset of features
 - for each character in the table
 - find its k nearest neighbors
 - classify it
 - update confusion matrix
- Evaluate confusion matrix
- Choose best subset of features

C	f_1	f_2	f_3	f_4

I									
	•								
S	•								
a	q	0	20	10	35				
w	p	1	12	55	10				
	b	3	50	10	8				
	a	65	1	2	0				
		a	b	p	q	•	•		

And called it



Editing the database - eliminating unnecessary elements

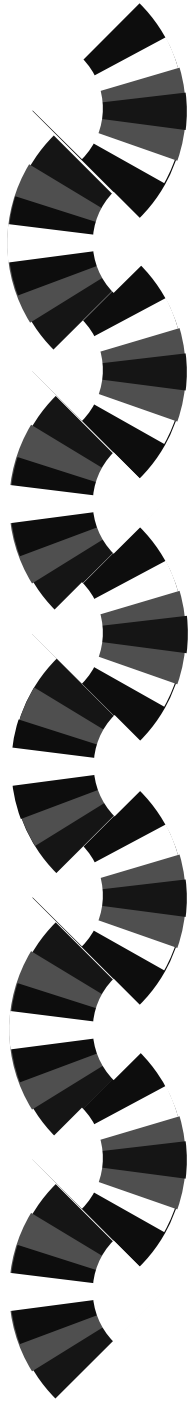
For each element in the ORIGINAL
training set

find its k-nearest neighbors in
reduced table

classify and update confusion
matrix

C	f ₁	f ₂	f ₃	f ₄

•					
•					
q	0	10	10	39	
p	1	14	50	12	
b	5	40	10	8	
a	58	1	2	0	
	a	b	p	q	• •



The final experiment

- ▶ Get the test image
 - Threshold it
 - Apply noise cleaning (if useful)
 - Compute its connected components
 - Build the table of features
 - Classify each component using the edited training set
 - Compute the confusion matrix
 - Write your report