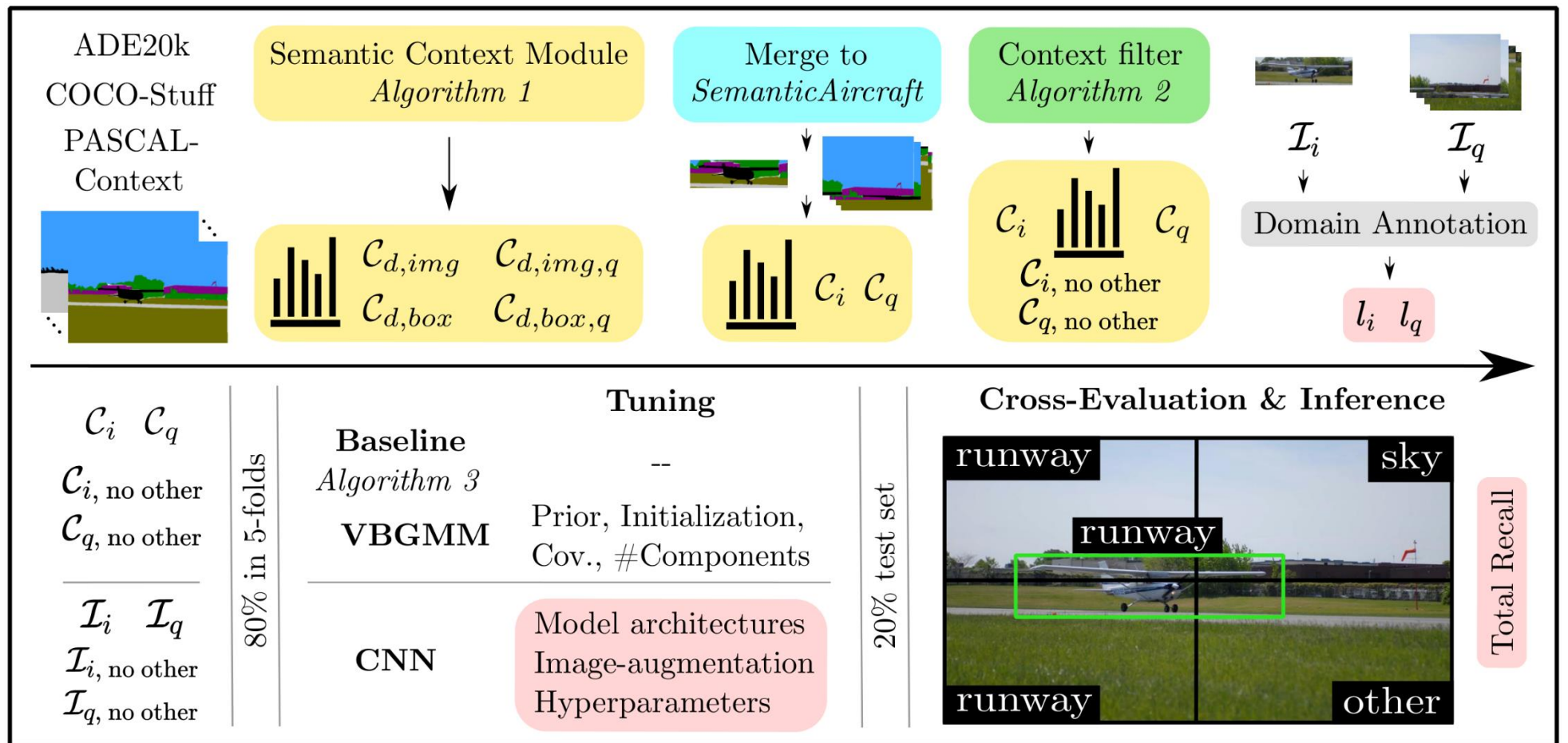


# Visual Semantic Context Encoding for Aerial Data Introspection and Domain Prediction

KRIEGLER, Andreas (Master thesis), STEININGER, Daniel and WÖBER, Wilfried



## Motivation

Semantic context only learned implicitly  
Domain prediction for downstream tracking

## Contributions

Simple context extraction for explicit priors  
*SemanticAircraft*: Harvested & filtered  
Domain prediction via novel baseline,  
neural network & mixture model  
Data introspection gives important insights

### Algorithm 3 Thresholding domain prediction

**Requisites:** Set of context vectors  $C_{i/q}$ . Set of domains  $\mathcal{D}$  and for every domain and superclass  $s$  consisting of classes  $c$  a certain range  $r_{x,y}$  and weight  $w_{x,y}$ . Domain-prediction threshold of  $th$  and a decrease  $th_d$ .

```

1: function TDP( $C_{i/q}, \mathcal{D}, \mathcal{S}, \mathcal{R}, \mathcal{W}, th, th_d$ )
2:   for  $c$  in  $\mathcal{C}$  do
3:      $d.s \leftarrow \mathbf{0} : \mathbf{0} \in \mathbb{R}^{n \times 1}$ 
4:     for  $d$  in  $\mathcal{D}$  do
5:       for  $s$  in  $\mathcal{S}$  do
6:          $s.s \leftarrow \sum_i c_i, \forall c \in s$ 
7:         if  $s.s \in [r_{d,s,l}, r_{d,s,u}]$  then
8:            $d.s_d \leftarrow d.s_d + w_{d,s}$ 
9:         if  $\max(d.s) > th$  then
10:           $l_c \leftarrow \text{argmax}(d.s)$ 
11:        else
12:           $th \leftarrow th - th_d$ 
13:   return  $l$ 

```

$\triangleright$  For every context vector  
 $\triangleright$  Initialize the domain score  
 $\triangleright$  And for every dataset  
 $\triangleright$  And superclass in that dataset  
 $\triangleright$  Aggregate context of all classes  
 $\triangleright$  Check if score is in range  
 $\triangleright$  Add a weight to the domain score  
 $\triangleright$  Take the top-1 domain  
 $\triangleright$  And assign the domain label  
 $\triangleright$  Or decrease threshold until domain is found  
 $\triangleright$  Return domain labels for every image patch

Table 1: Visual percentage-wise context for the *SemanticAircraft* dataset showing dominant sky-context. Context across all four quadrants was merged.

	building	elevation	object	pavement	person	plant	sky	soil	vehicle	waterbody
Instances	7.5	3.2	1.5	15.8	1.2	5.1	57.2	6.3	0.9	1.3
Quadrants	4.2	2.8	1.2	17.4	1.1	4.0	58.6	7.8	1.0	1.9

Table 2: Accuracy of all three models predicting domains of airplane instances and quadrants from *SemanticAircraft*.

	Instances		Quadrants	
	Including	Other Excluding	Including	Other Excluding
Baseline	0.588 ± 0.015	0.796 ± 0.011	0.639 ± 0.017	<b>0.799 ± 0.006</b>
VBGMM	0.586 ± 0.048	0.712 ± 0.06	0.539 ± 0.029	0.637 ± 0.083
ResNet18	<b>0.716 ± 0.015</b>	<b>0.854 ± 0.011</b>	<b>0.692 ± 0.013</b>	0.778 ± 0.006

