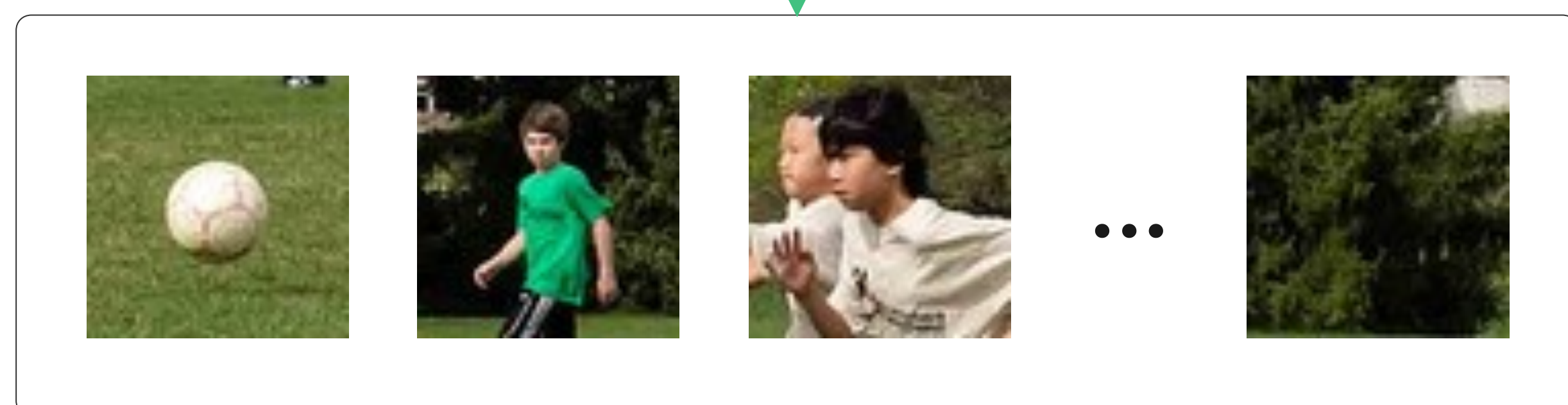




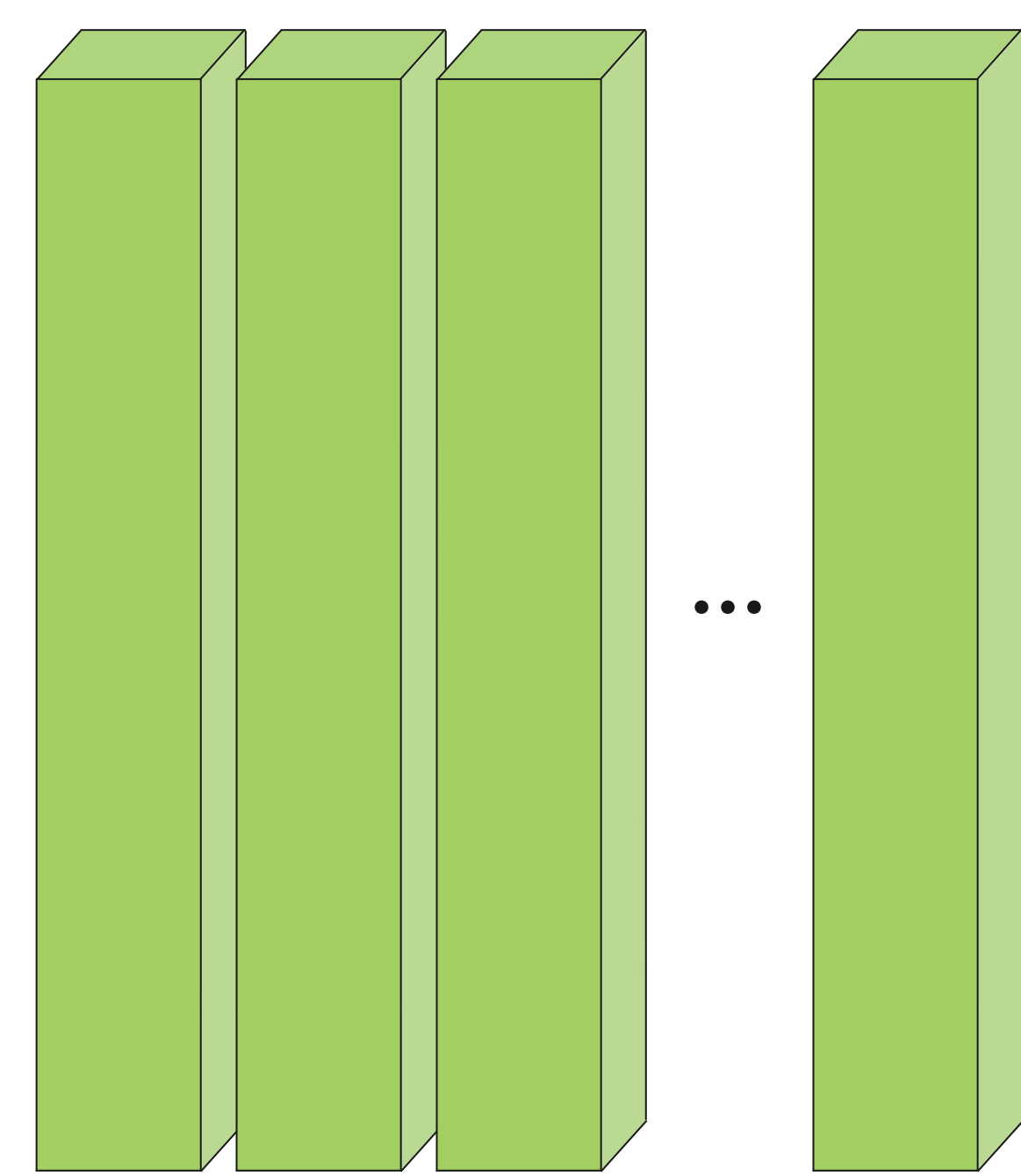
**Faster  
R-CNN**



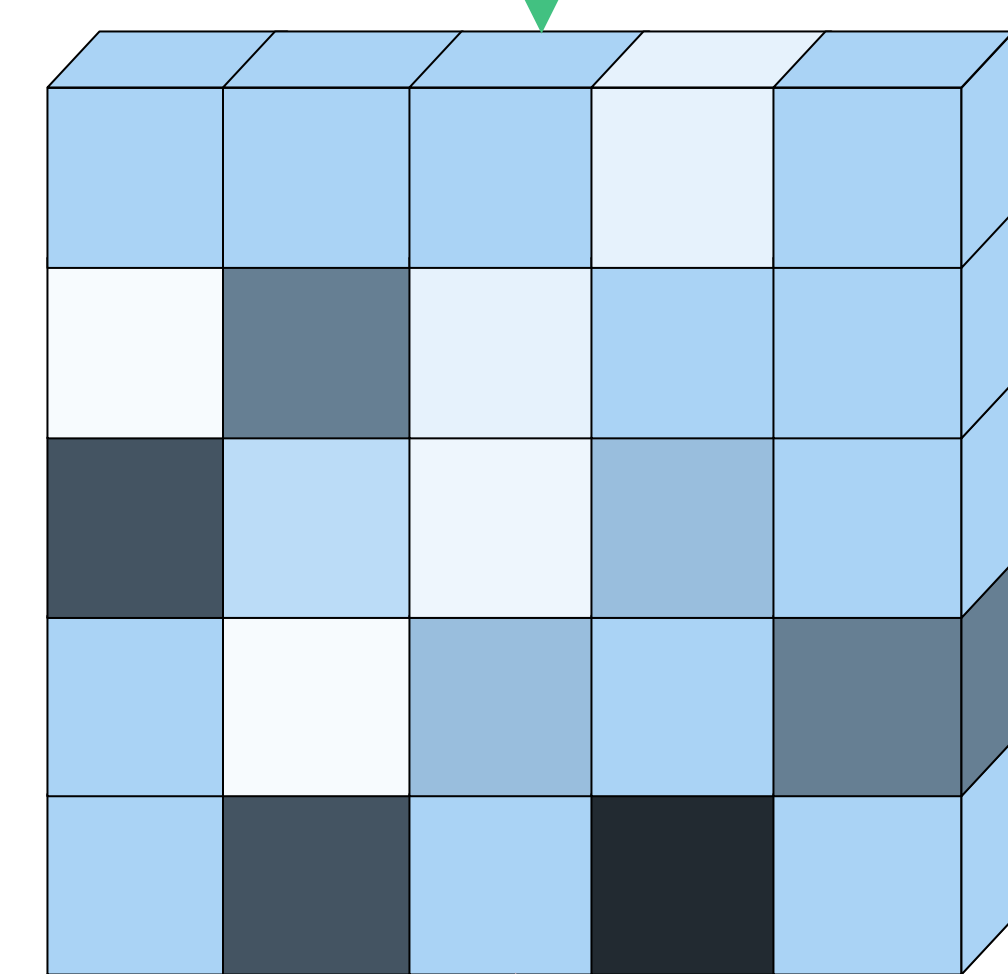
children wearing team uniforms , playing soccer in a grassy field .

**RNN**

$\{f_i[1], f_i[2], \dots, f_i[k]\}$



$S(f_i, f_c)$   
 $S(f'_i, f_c)$   
 $S(f_i, f'_c)$



**Positive  
Attention**

**Negative  
Attention**

**Aggregation**

$L_{fine}$   $L_{coarse}$

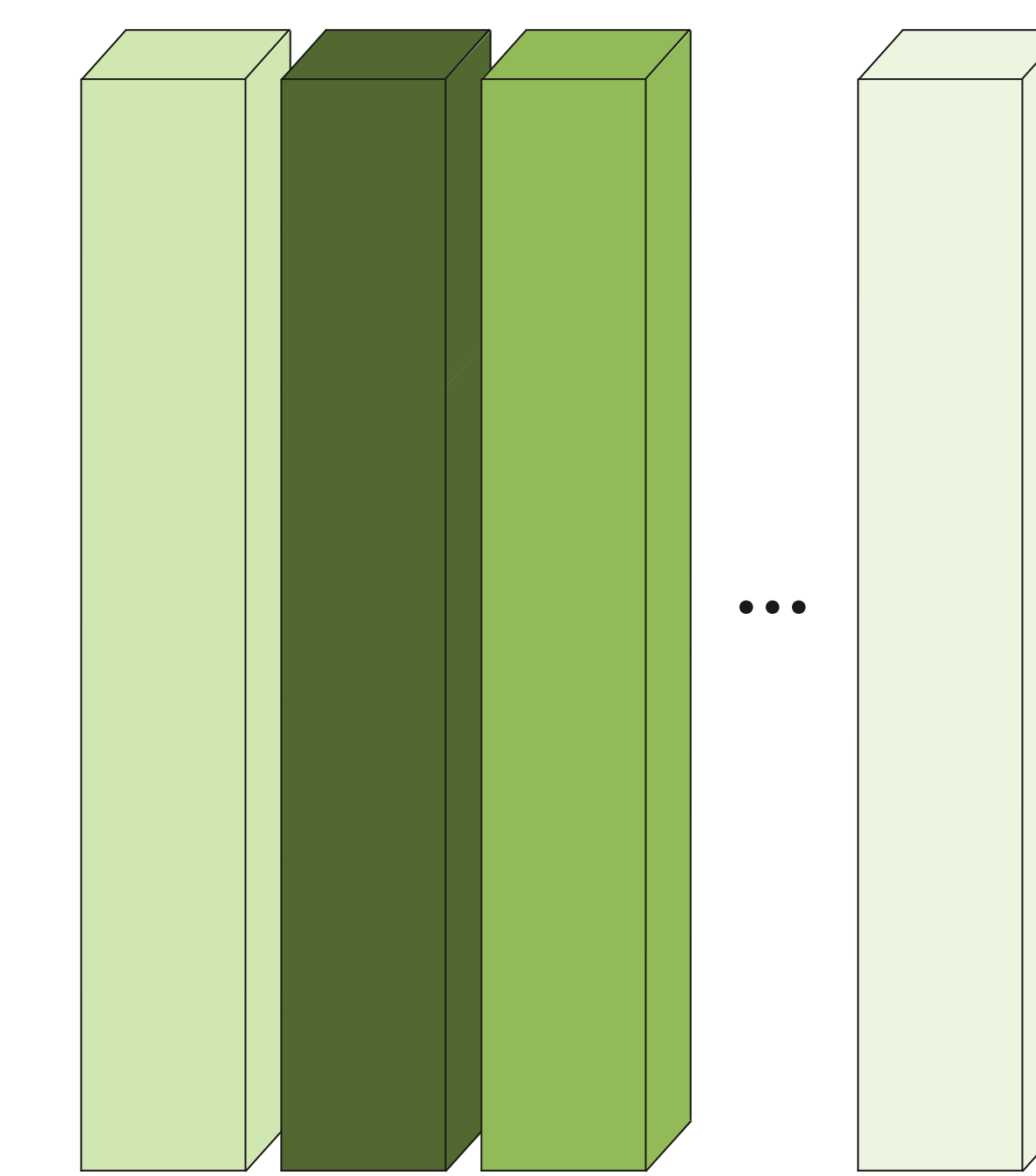
**Fine-grained  
Constraints**

$S^*(f_i^*, f_c^*)$

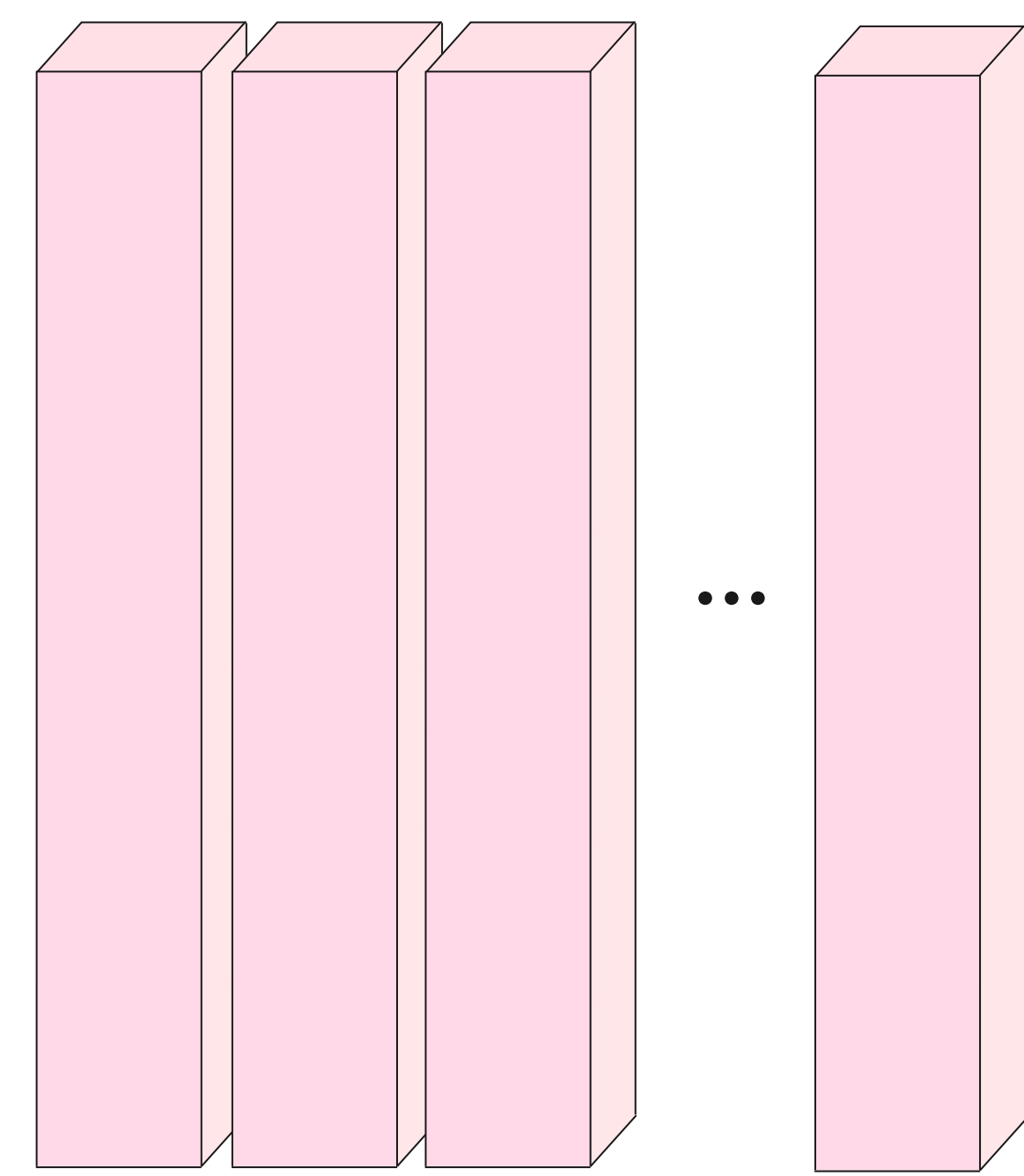
**Cosine  
Similarity**

**Adaptive  
Pooling**

$\{f_i^*[1], f_i^*[2], \dots, f_i^*[k]\}$



$\{f_c[1], f_c[2], \dots, f_c[l]\}$



**Aggregation**

$\{f_c^*[1], f_c^*[2], \dots, f_c^*[l]\}$

