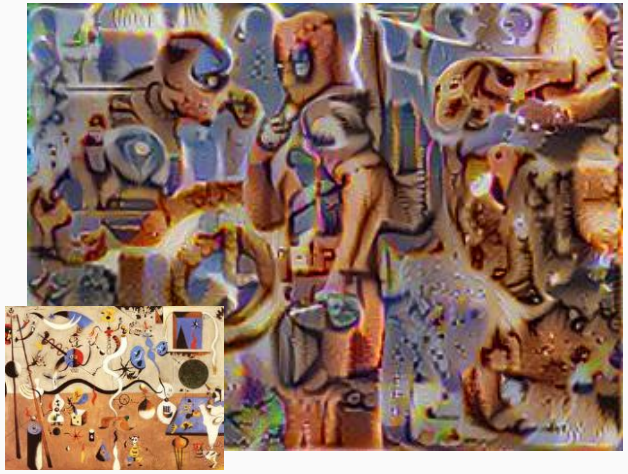
Style Transfer

Content Image ($I_{content}$)

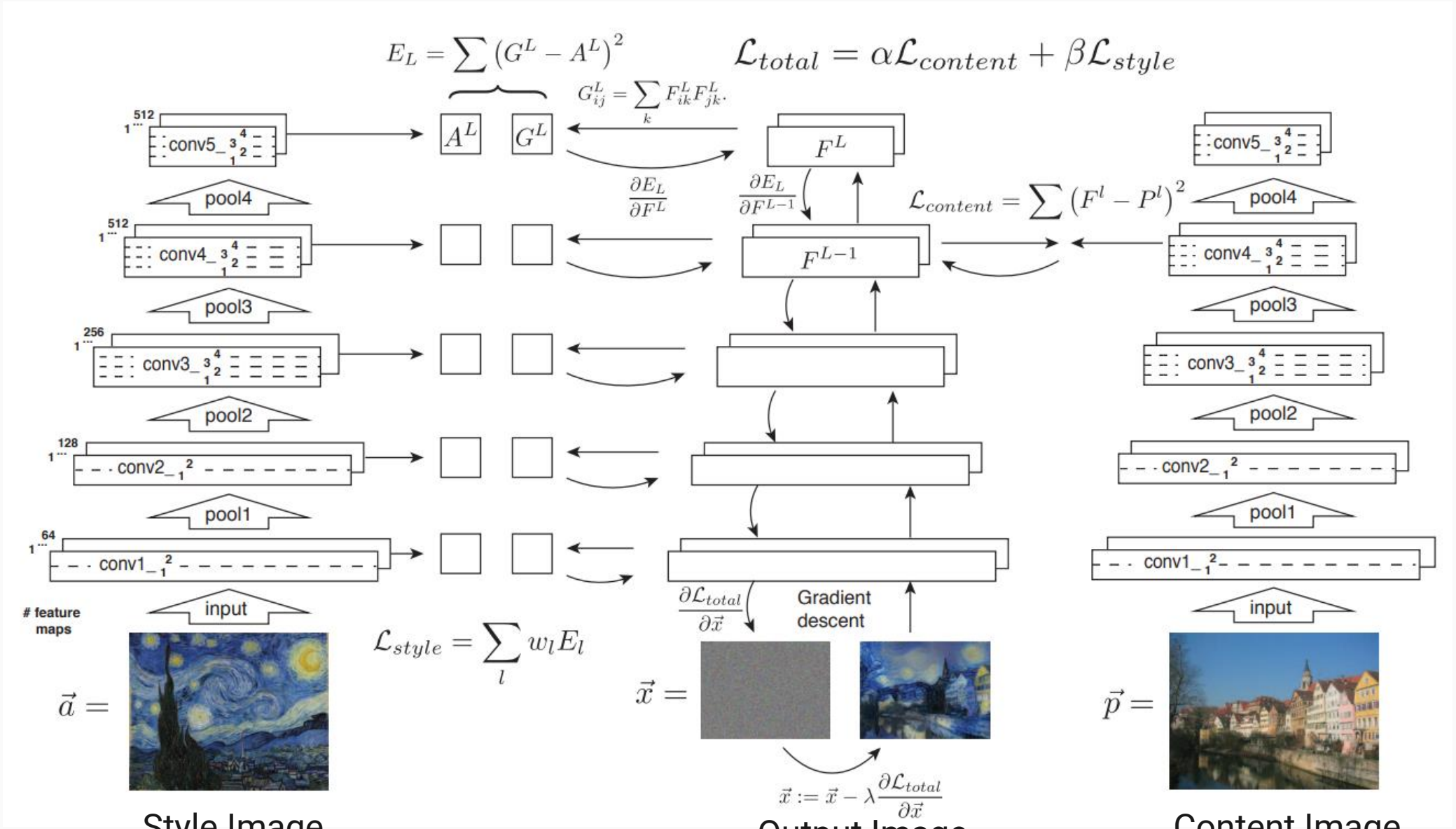
Style Image (I_{style})

Output Image (I_{output})

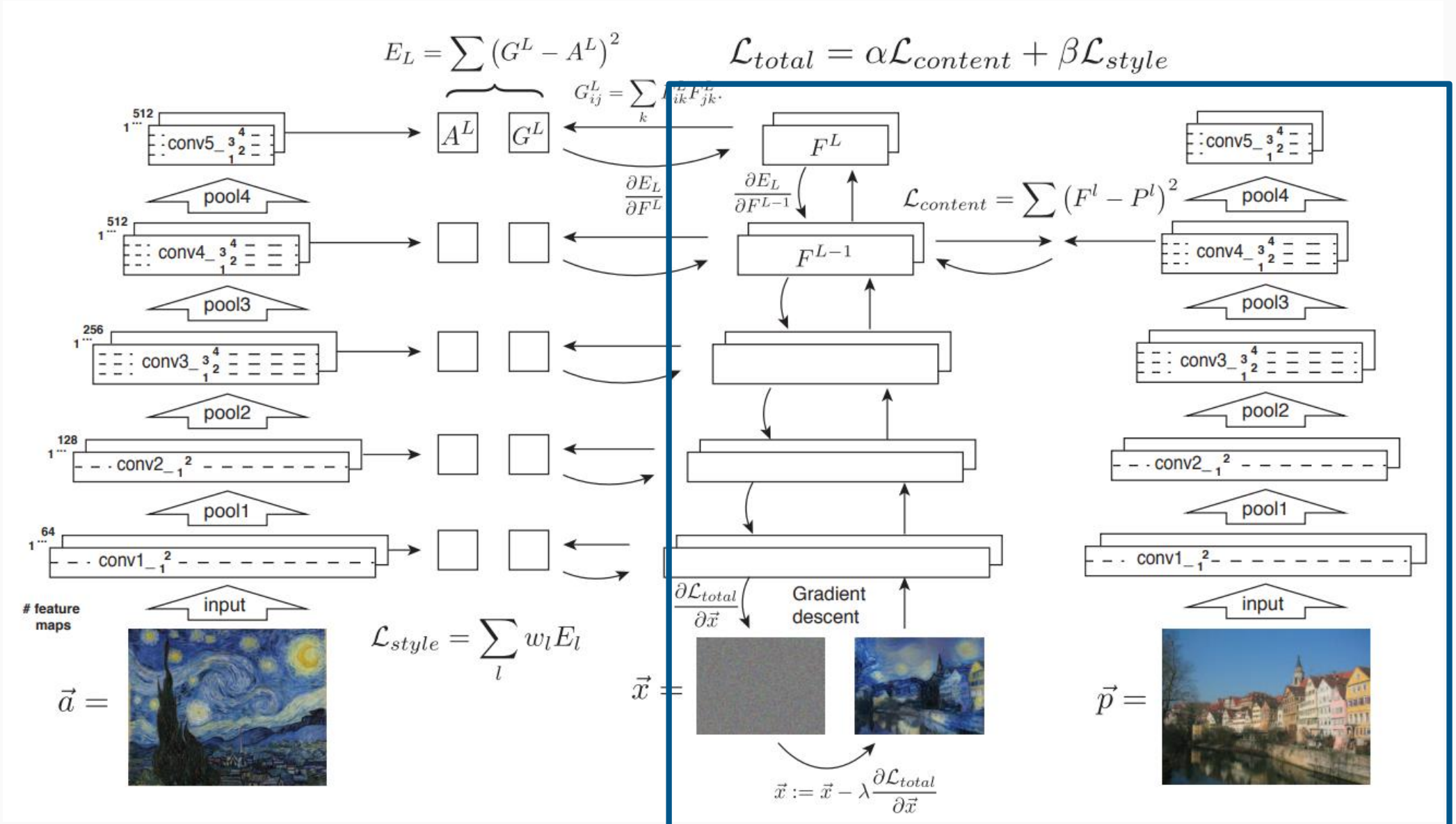
Style Transfer



Style Transfer



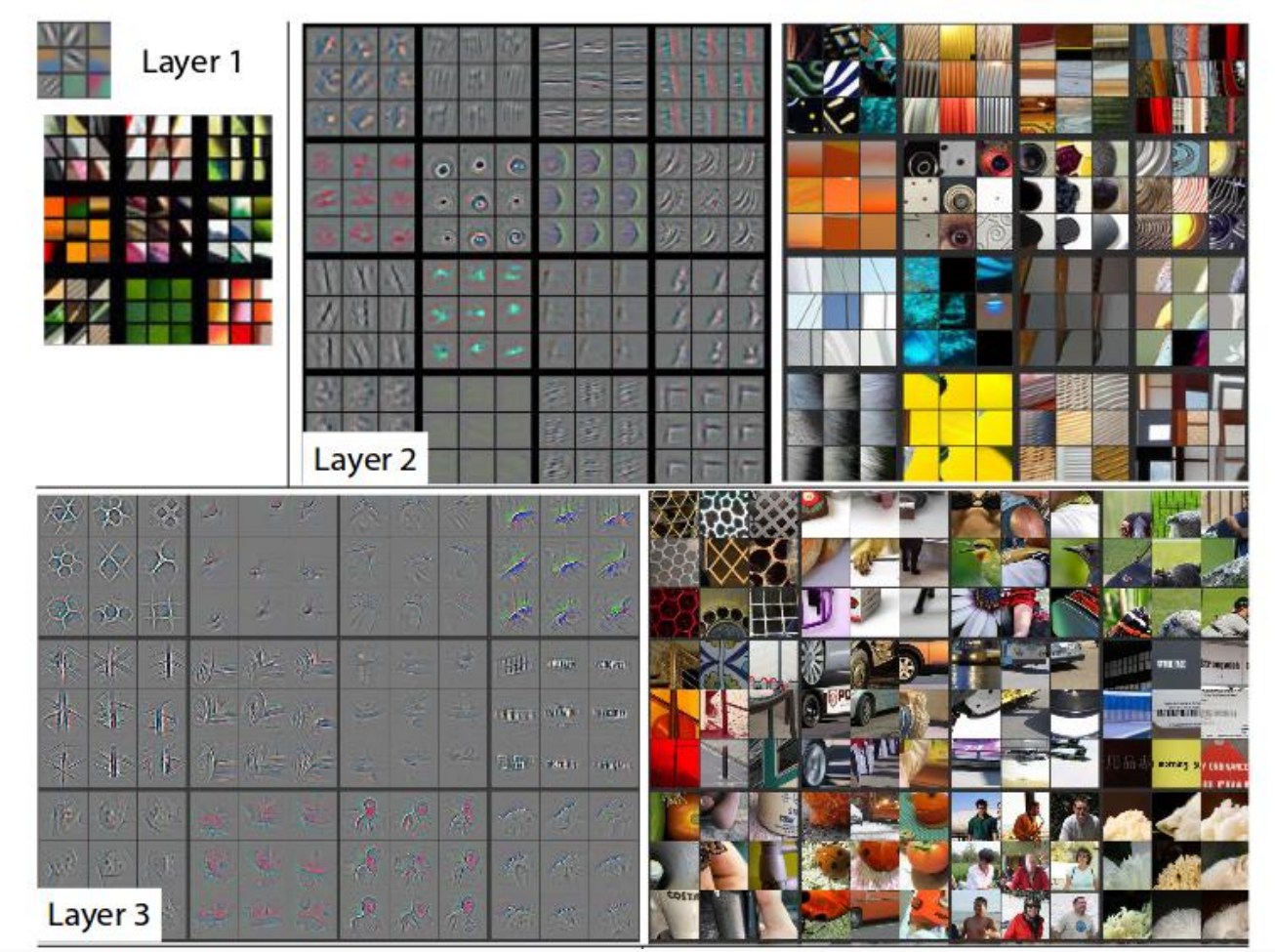
Style Transfer



1. How to represent content

Style Transfer

Content Representation



Style Transfer

Content Representation

For “Conv1_1”

1. # of filter:

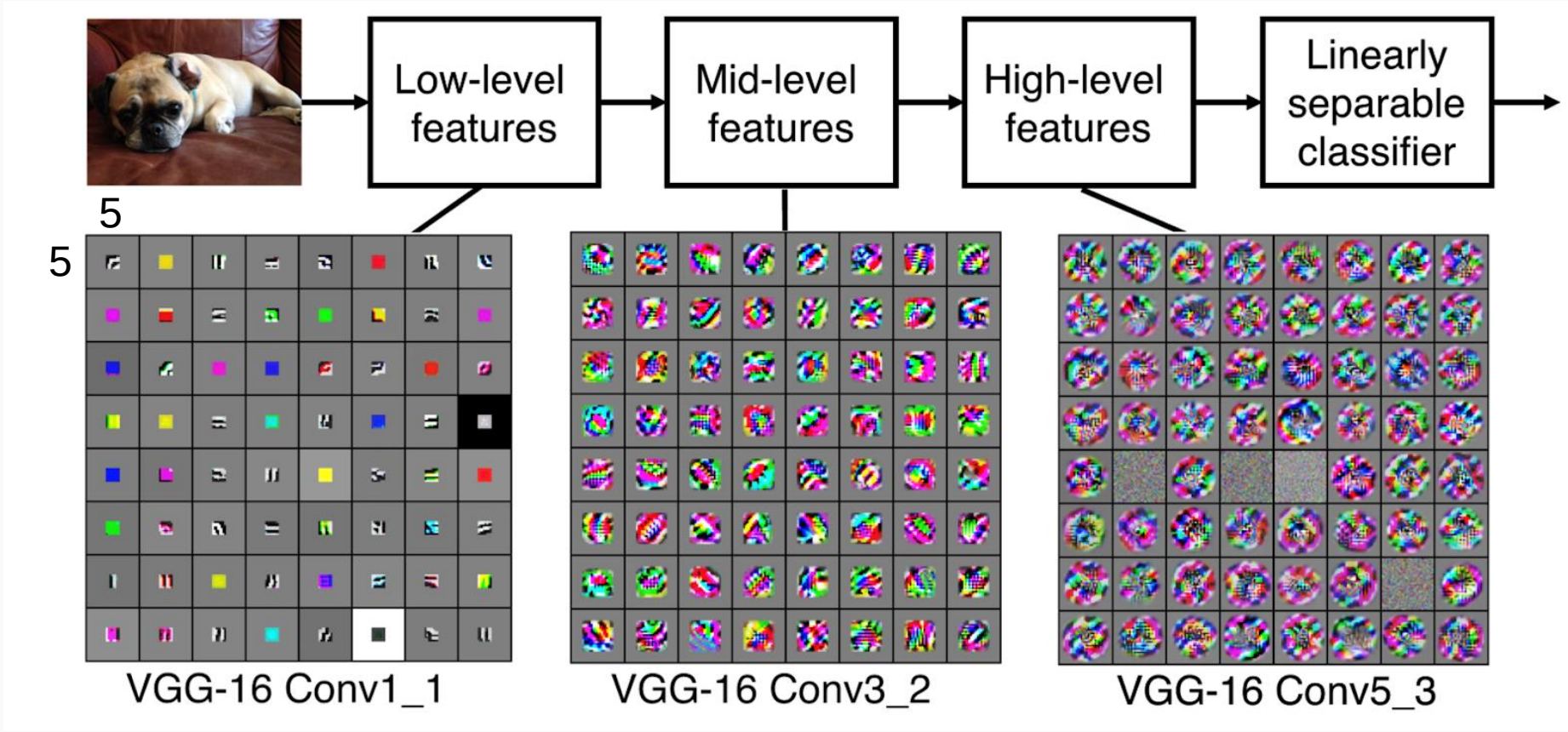
   ?

2. size of filter:

   ? x ?

3. # of feature:

   ? x ? x ?



Style Transfer

Content Representation

For “Conv1_1”

1. # of filter:

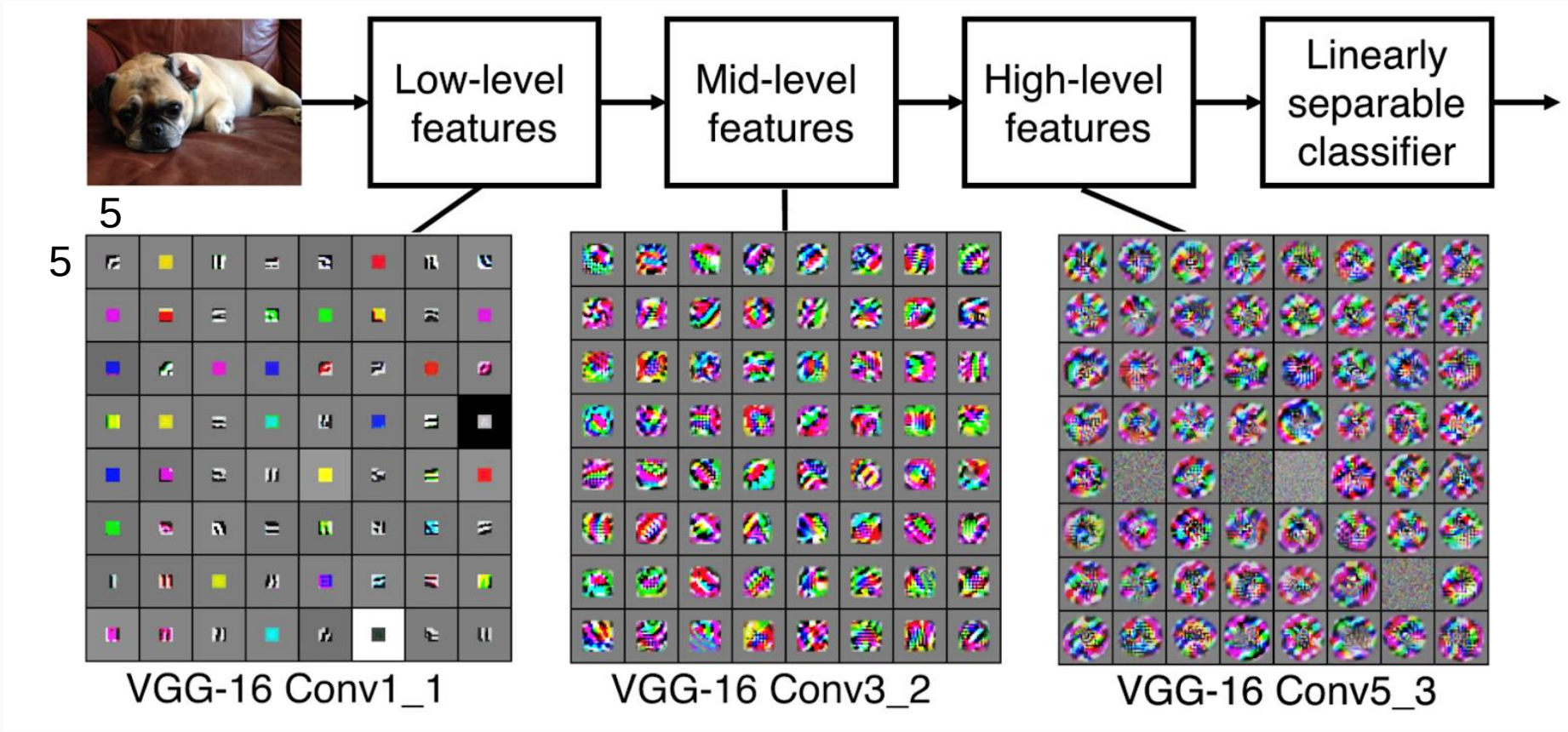
64

2. size of filter:

? x ?

3. # of feature:

? x ? x ?



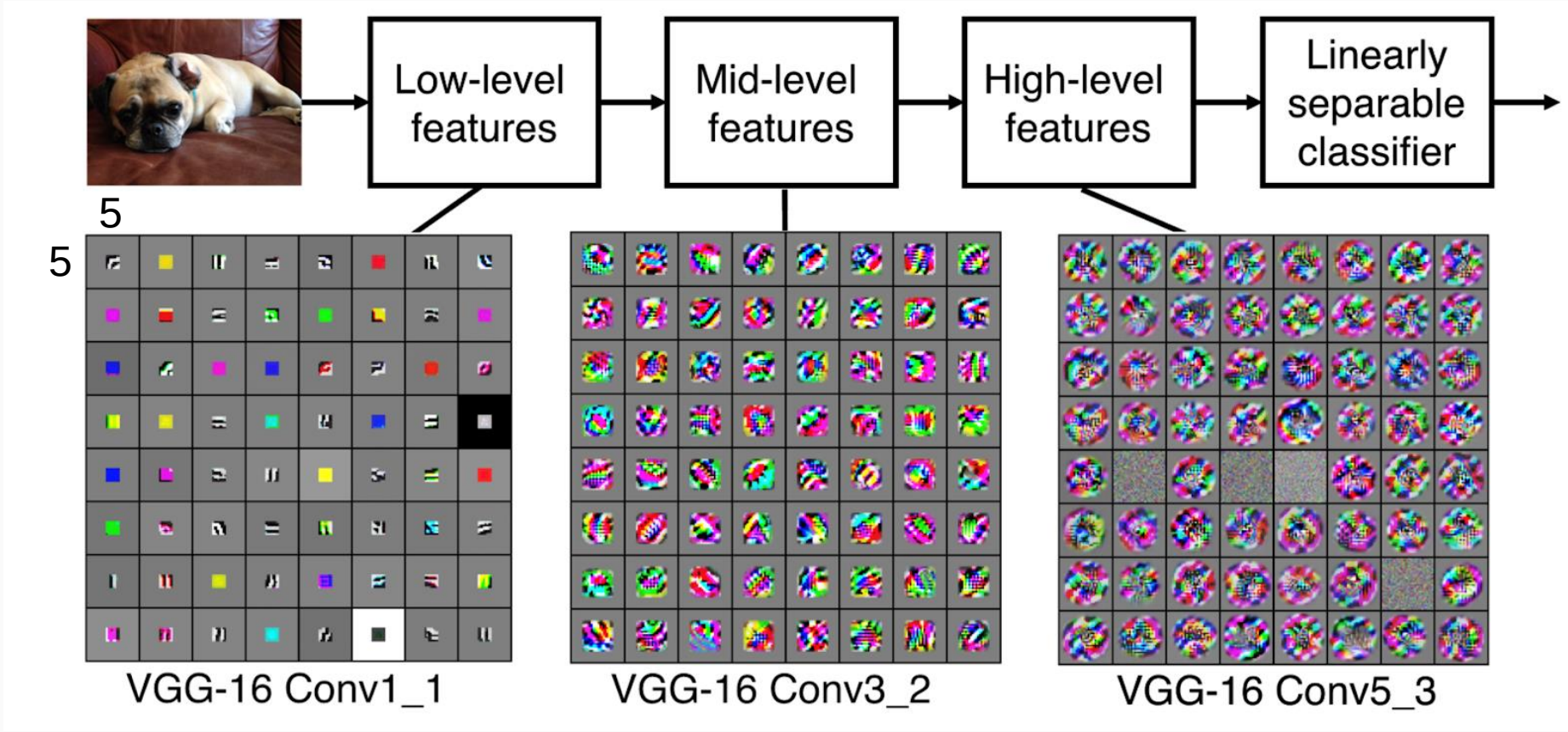
Content Representation

For “Conv1_1”

1. # of filter:
   64

2. size of filter:
   5 x 5

3. # of feature:
   ? x ? x ?



Style Transfer

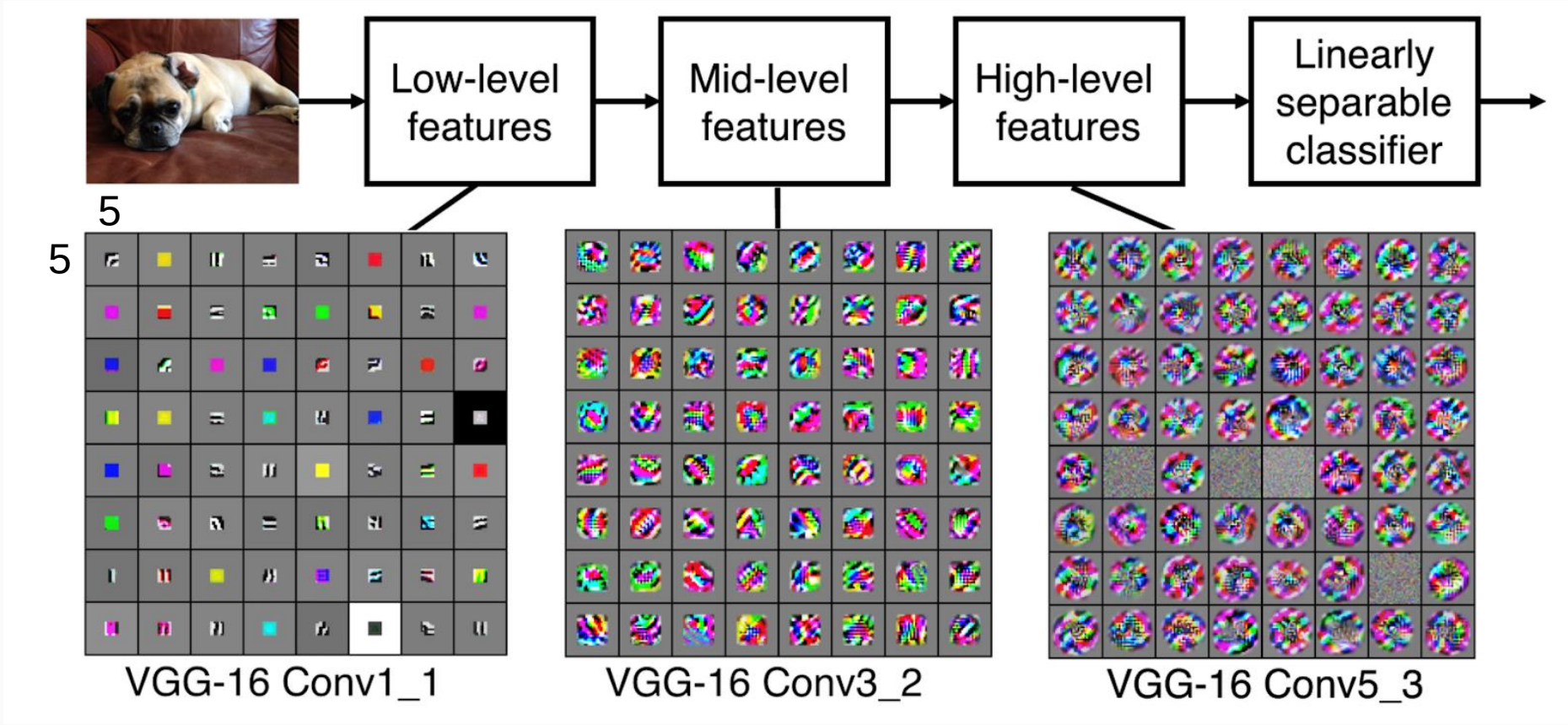
Content Representation

For “Conv1_1”

1. # of filter:
   64

2. size of filter:
   5 x 5

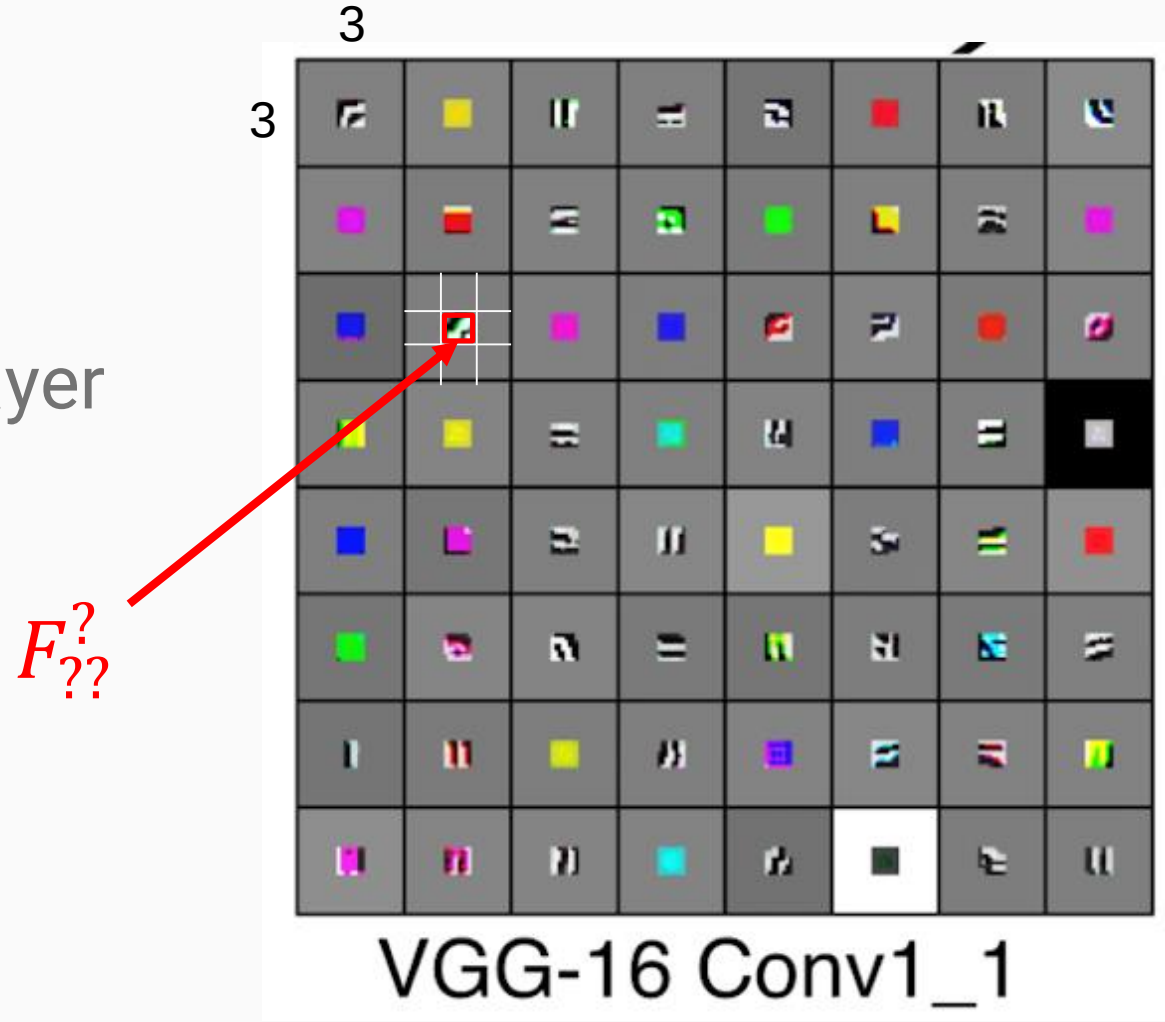
3. # of feature:
   64 x 5 x 5



Style Transfer

Content Representation

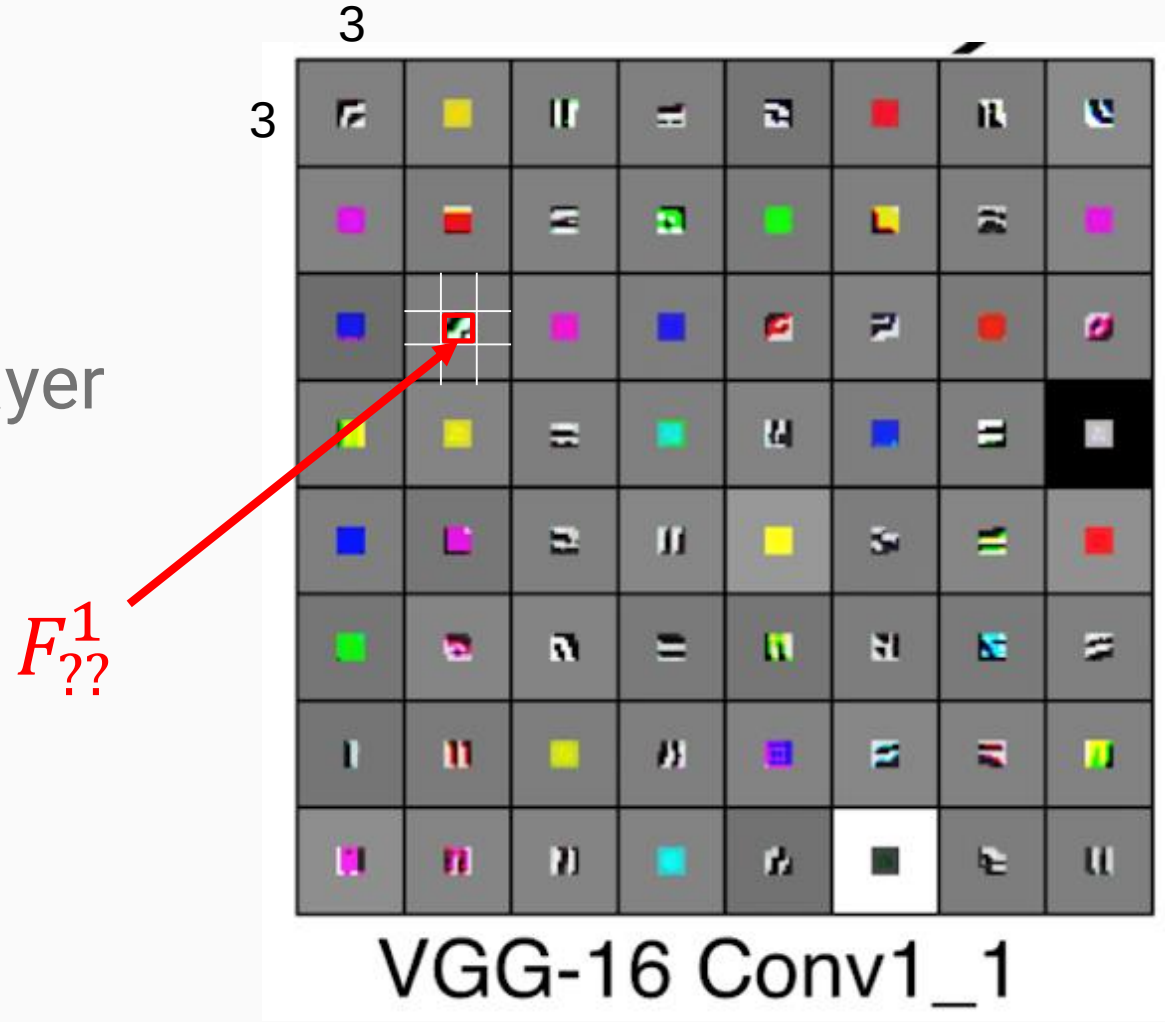
- l: index for layer
- i: index for filter
- F_{ij}^{l}: j^{th} feature in i^{th} filter of l^{th} layer



Style Transfer

Content Representation

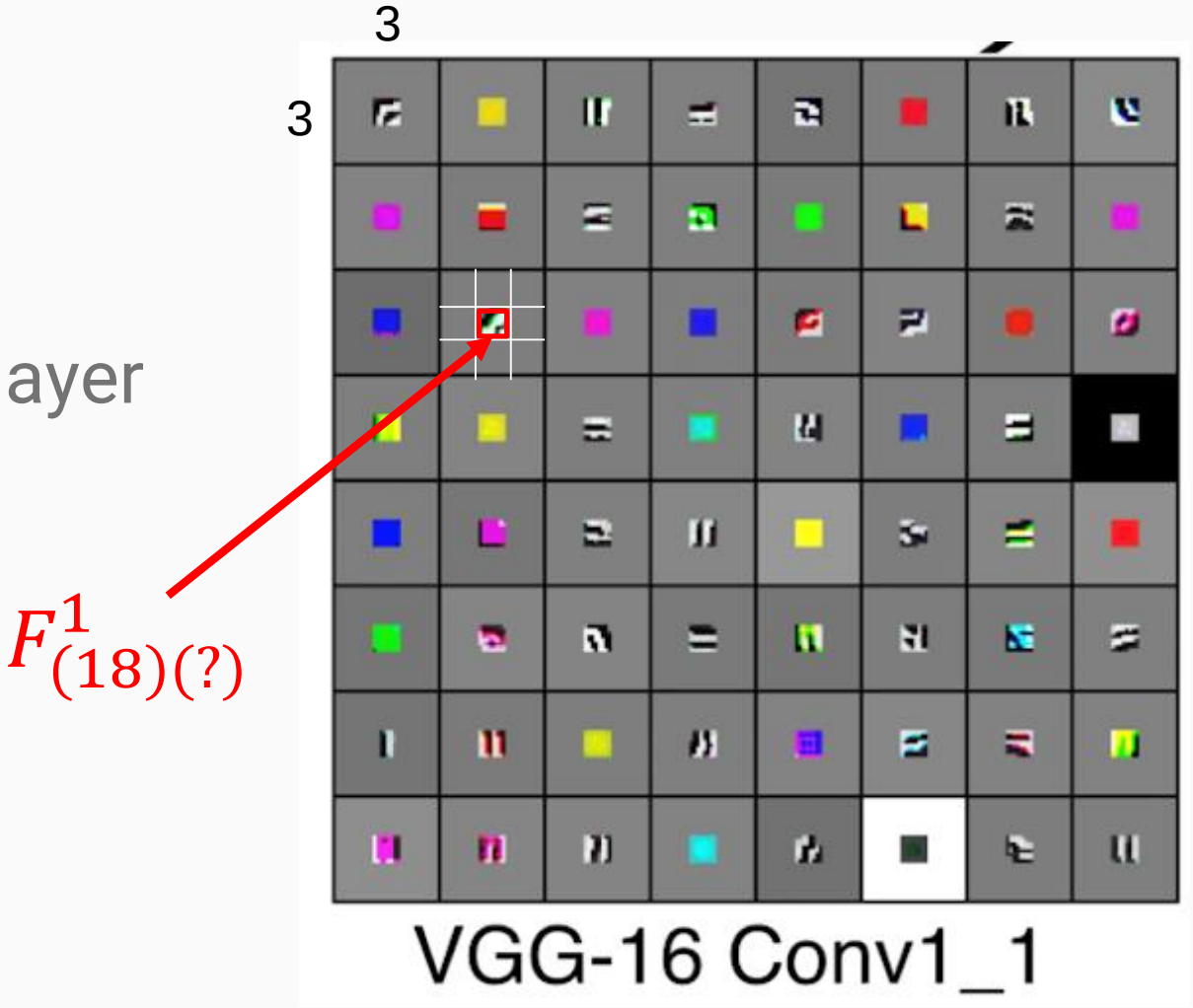
- l: index for layer
- i: index for filter
- F_{ij}^{l}: j^{th} feature in i^{th} filter of l^{th} layer



Style Transfer

Content Representation

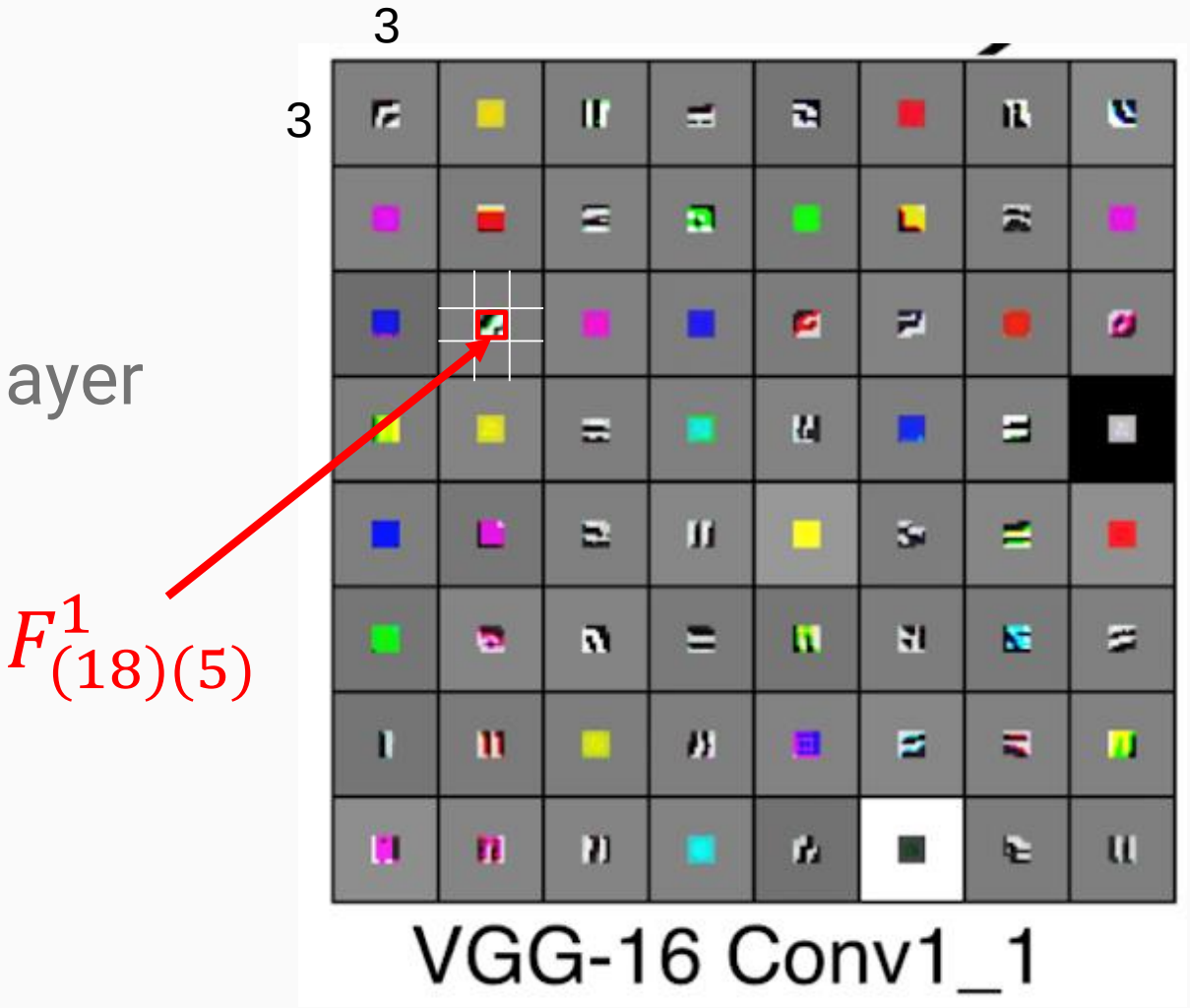
- l: index for layer
- i: index for filter
- F_{ij}^{l}: j^{th} feature in i^{th} filter of l^{th} layer



Style Transfer

Content Representation

- l: index for layer
- i: index for filter
- F_{ij}^{l}: j^{th} feature in i^{th} filter of l^{th} layer



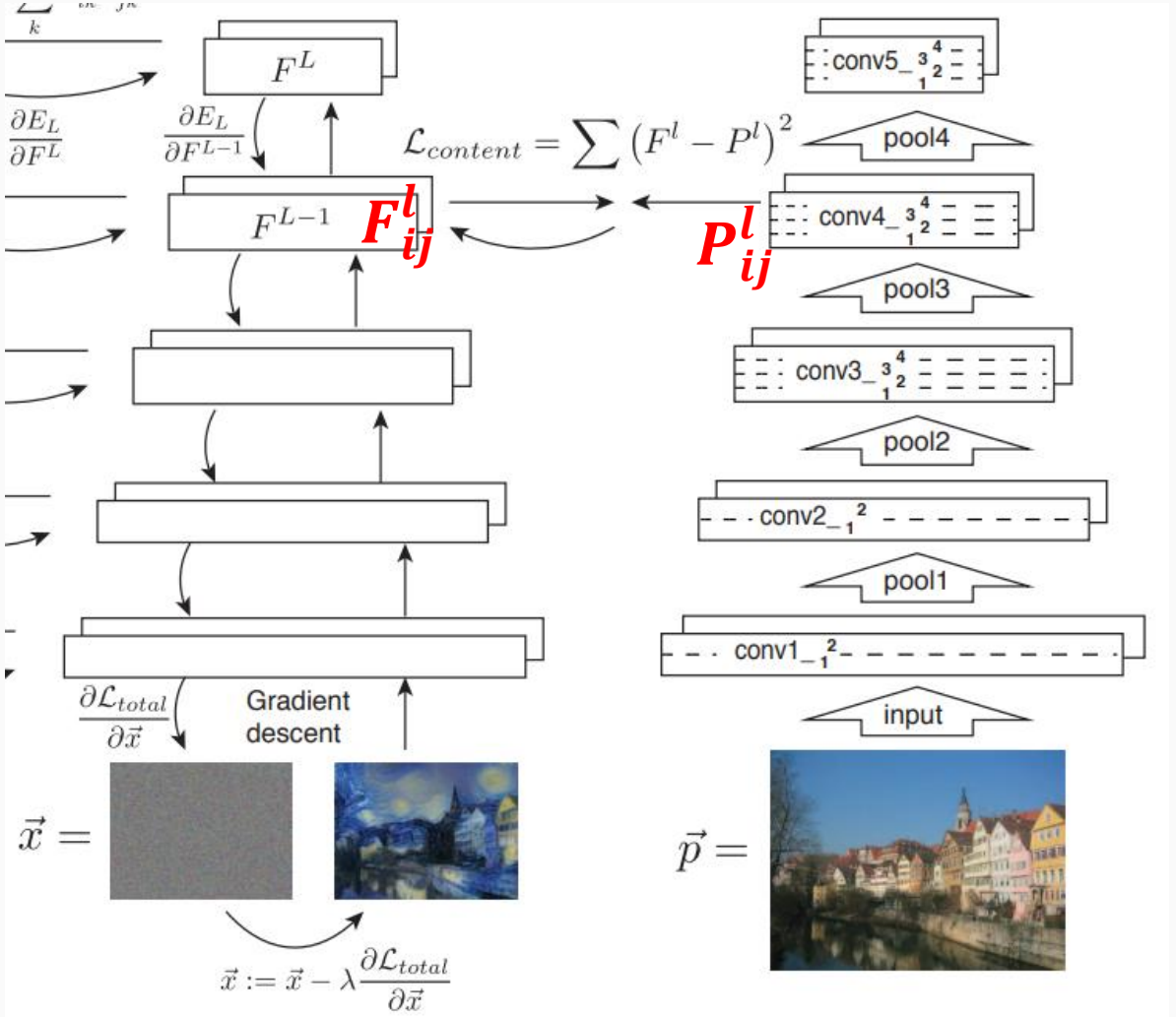
Style Transfer

Content Representation

- $L_{content}(\vec{p}, \vec{x}, l) =$

$$\frac{1}{2}\sum_{i,j}\left(F_{ij}^{l} - P_{ij}^{l}\right)^2$$

- Features (F_{i*}^{l}) for same area (j) should be similar.

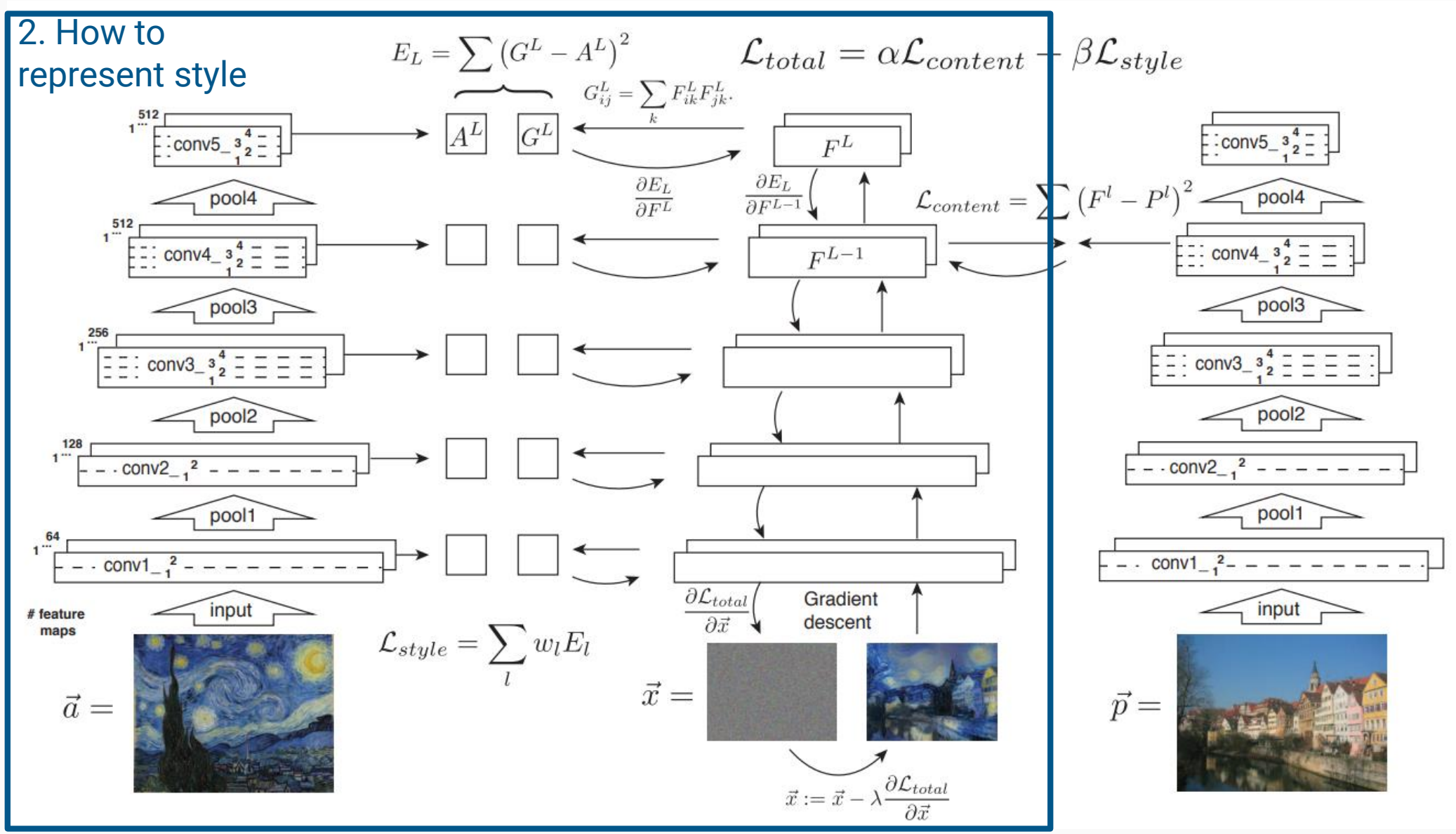


Output Image

Content Image

Style Transfer

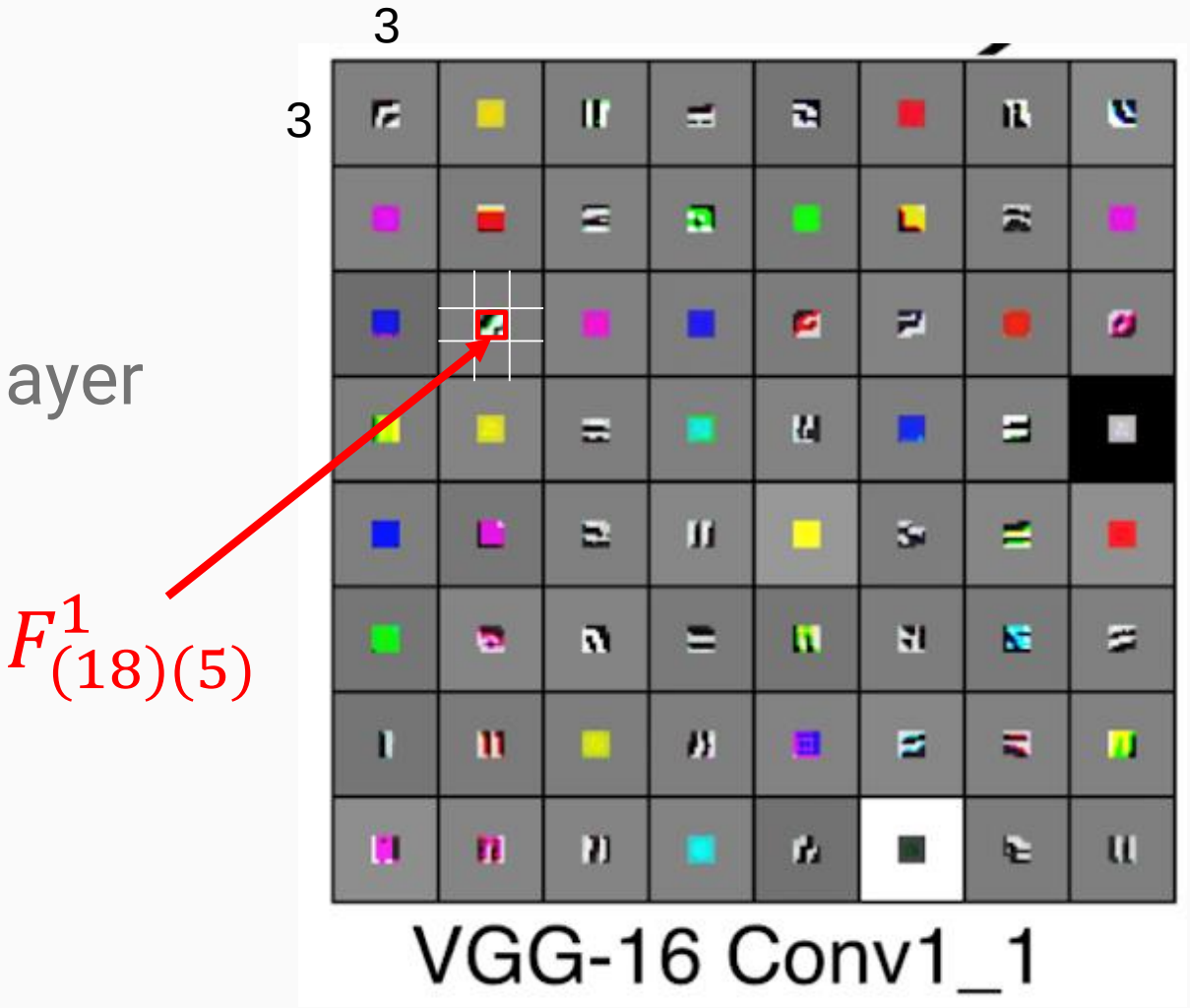
2. How to represent style



Style Transfer

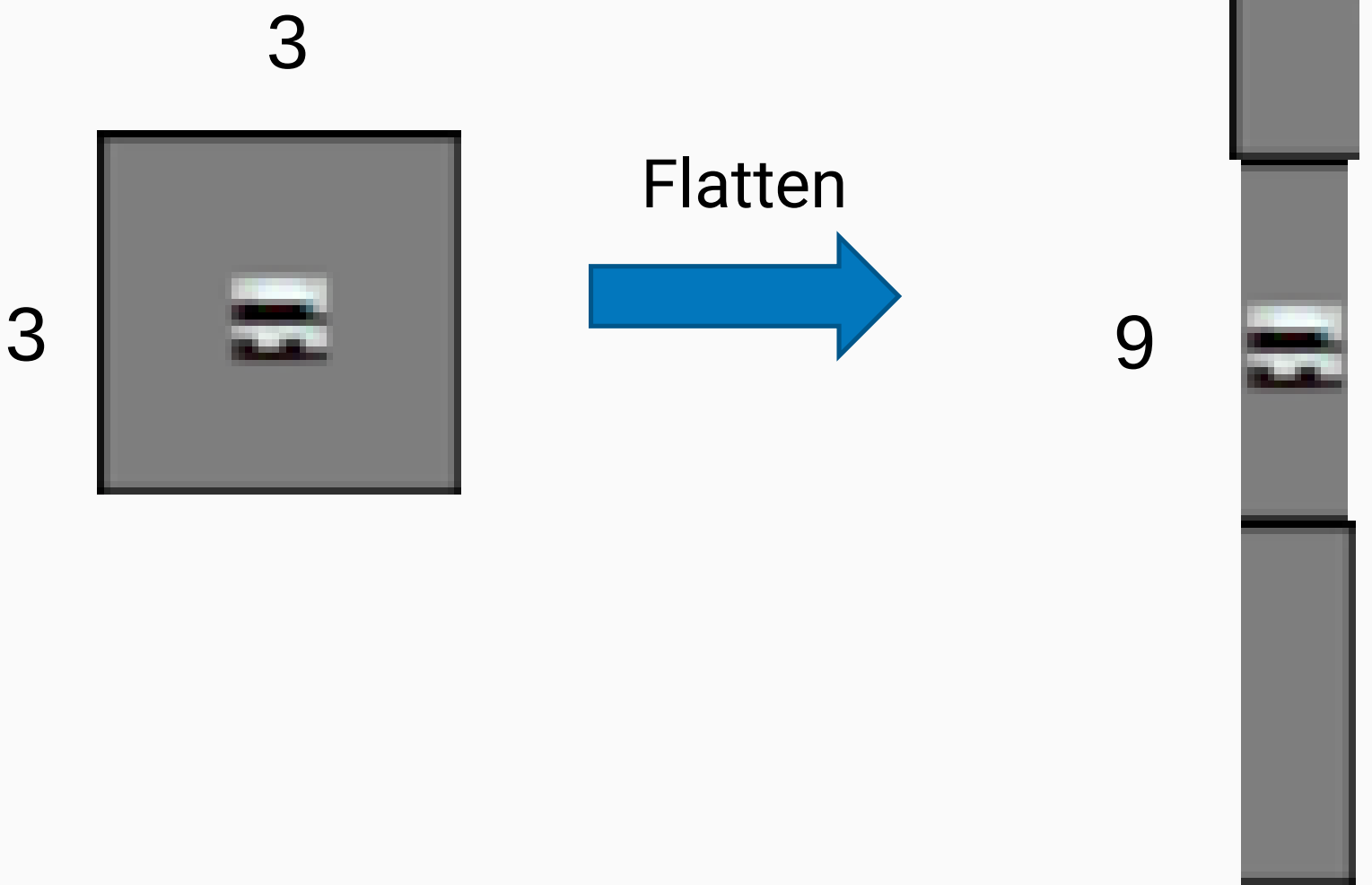
Style Representation

- l: index for layer
- i: index for filter
- F_{ij}^{l}: j^{th} feature in i^{th} filter of l^{th} layer



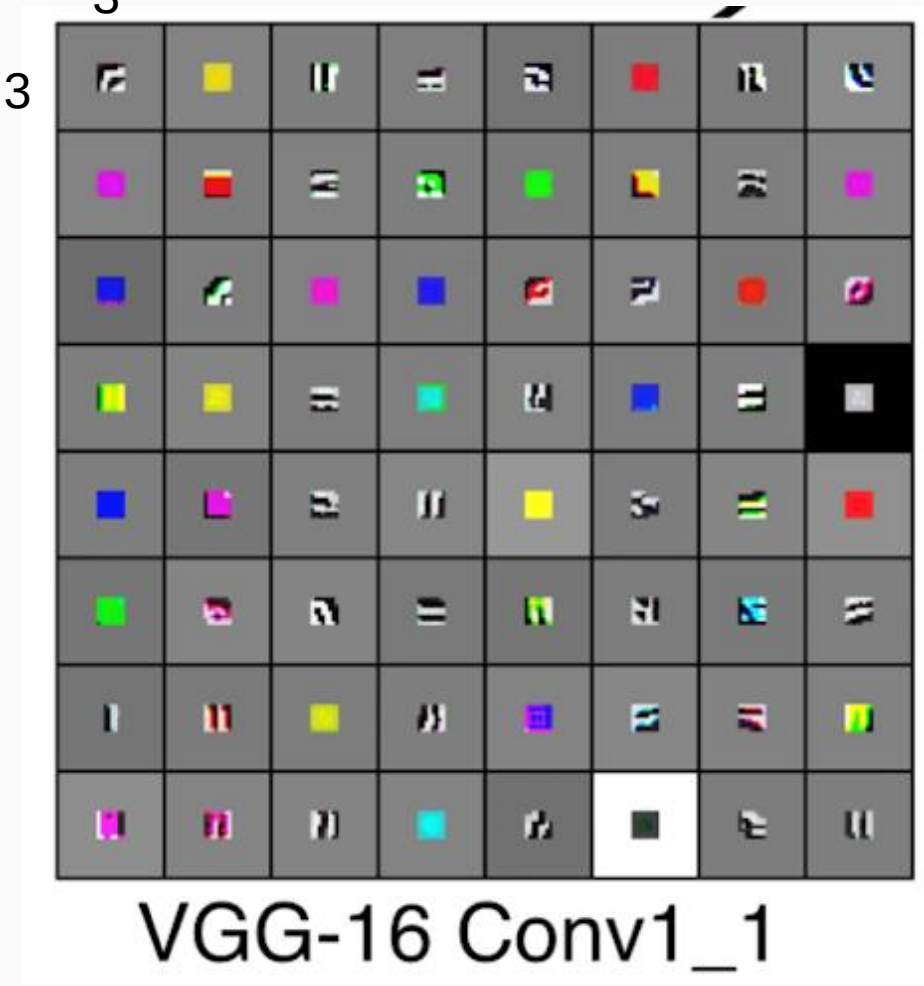
Style Transfer

Style Representation

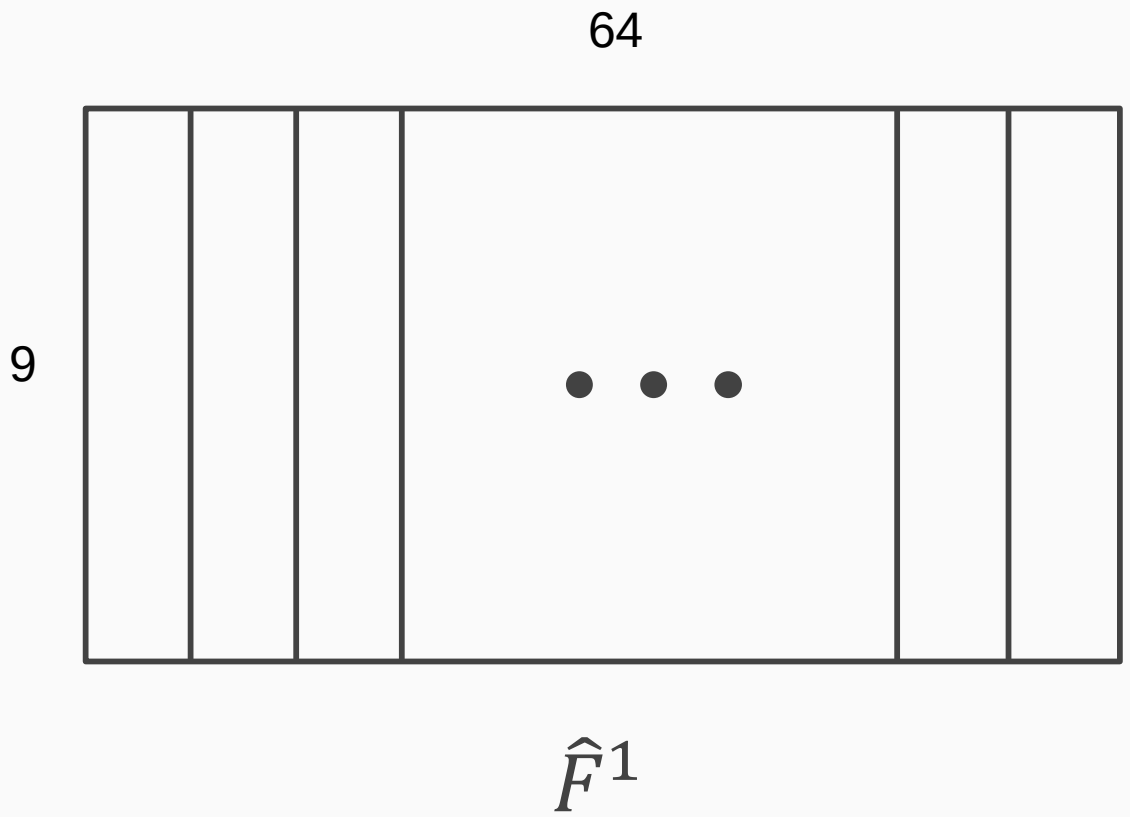


Style Transfer

Style Representation



Flatten



Style Transfer

Style Representation

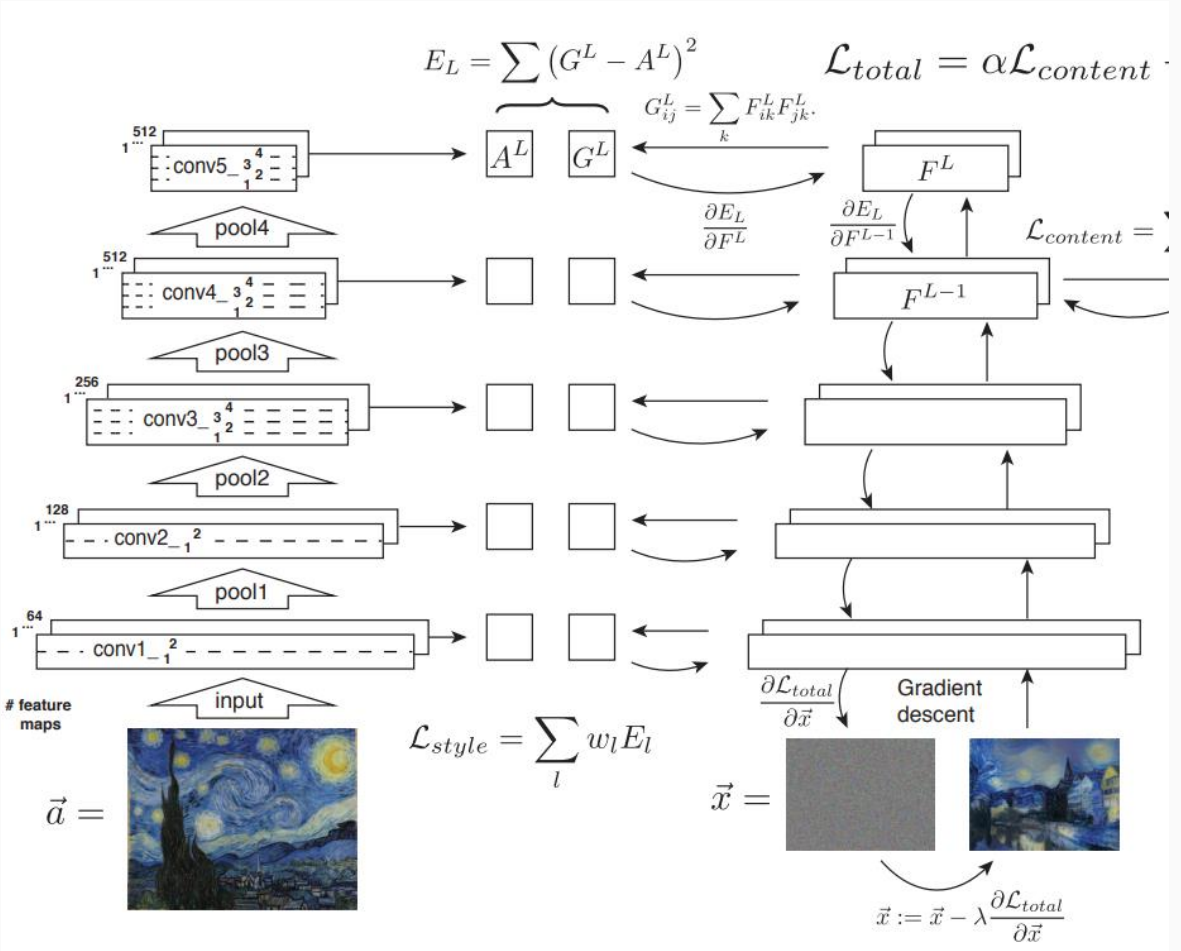
- $G^l = \left(\hat{F}^l\right)^T \hat{F}^l$

→ G^l: (# of filter) x (# of filter)

→ dot-product between features

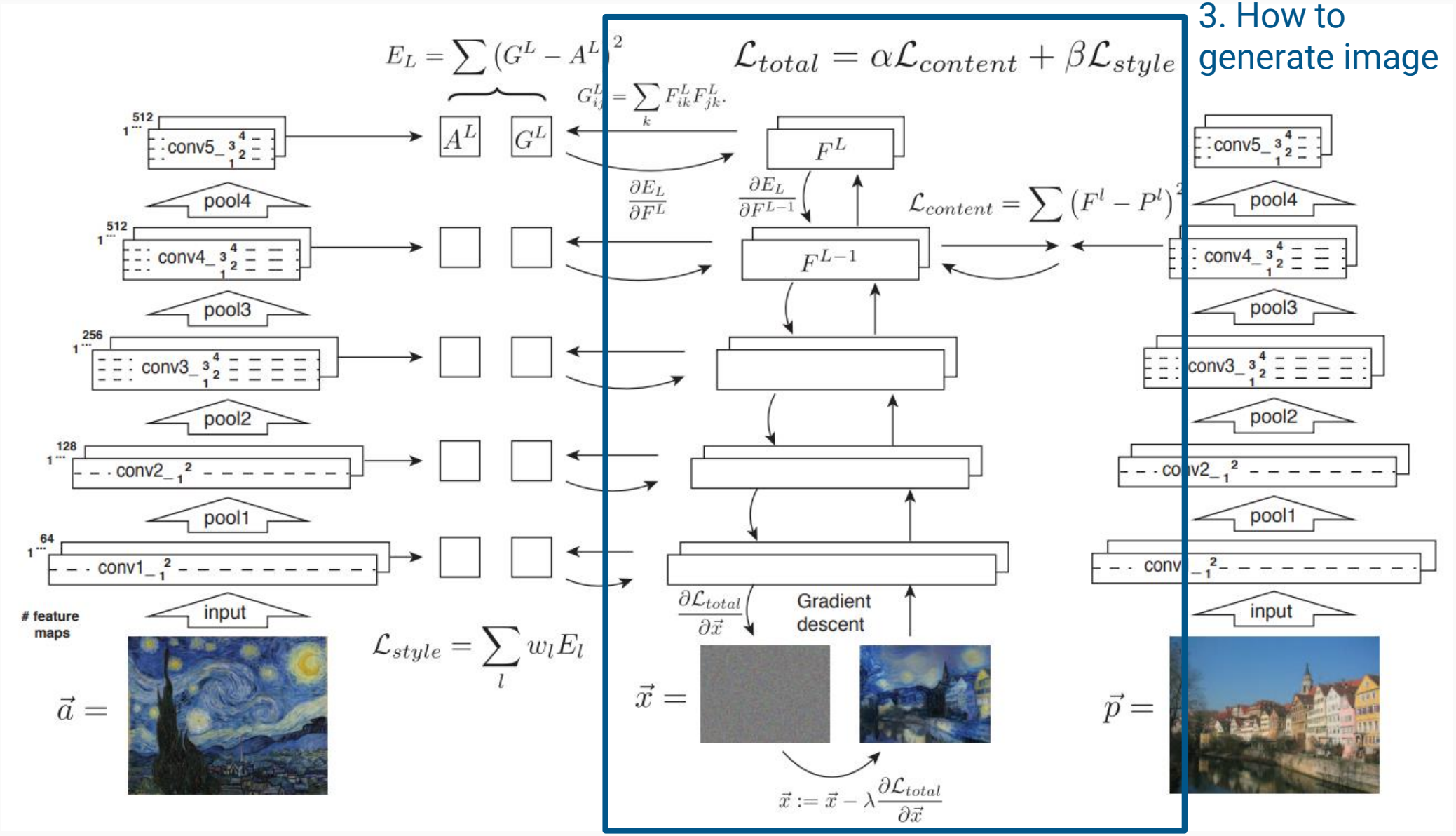
- $E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} \left(G^l - A^l\right)^2$
- $L_{style} = \sum_l w_l E_l$

- Relation between features (G^l) should be similar.



Style Transfer

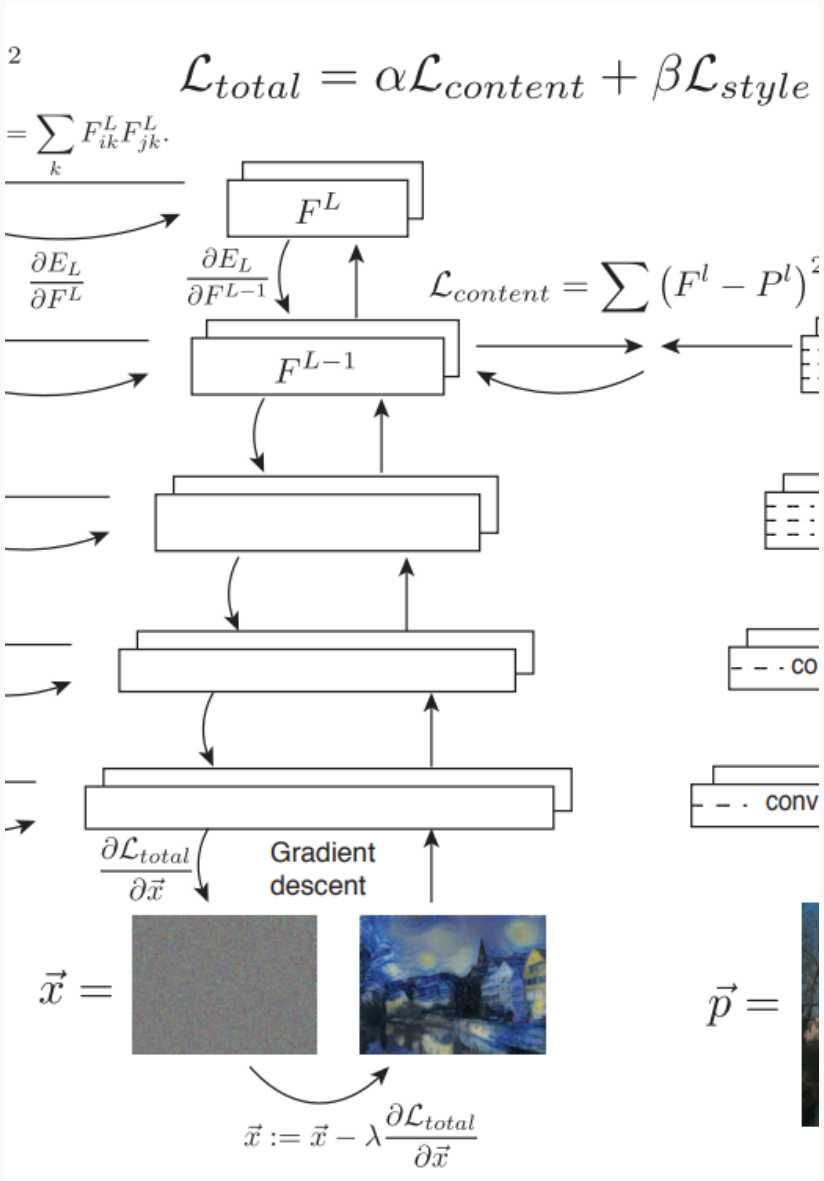
3. How to generate image



Style Transfer

Image generation

- $L_{total} = \alpha L_{content} + \beta L_{style}$
- Generate image via gradient descent $(\frac{dL_{total}}{dI_{output}})$
- No weight update (use pretrained VGG19 model)
- Update output image



Reference

[1] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
[2] https://web.stanford.edu/class/cs20si/2017/index.html
[3] https://github.com/Hvass-Labs/TensorFlow-Tutorials/blob/master/15_Style_Transfer.ipynb