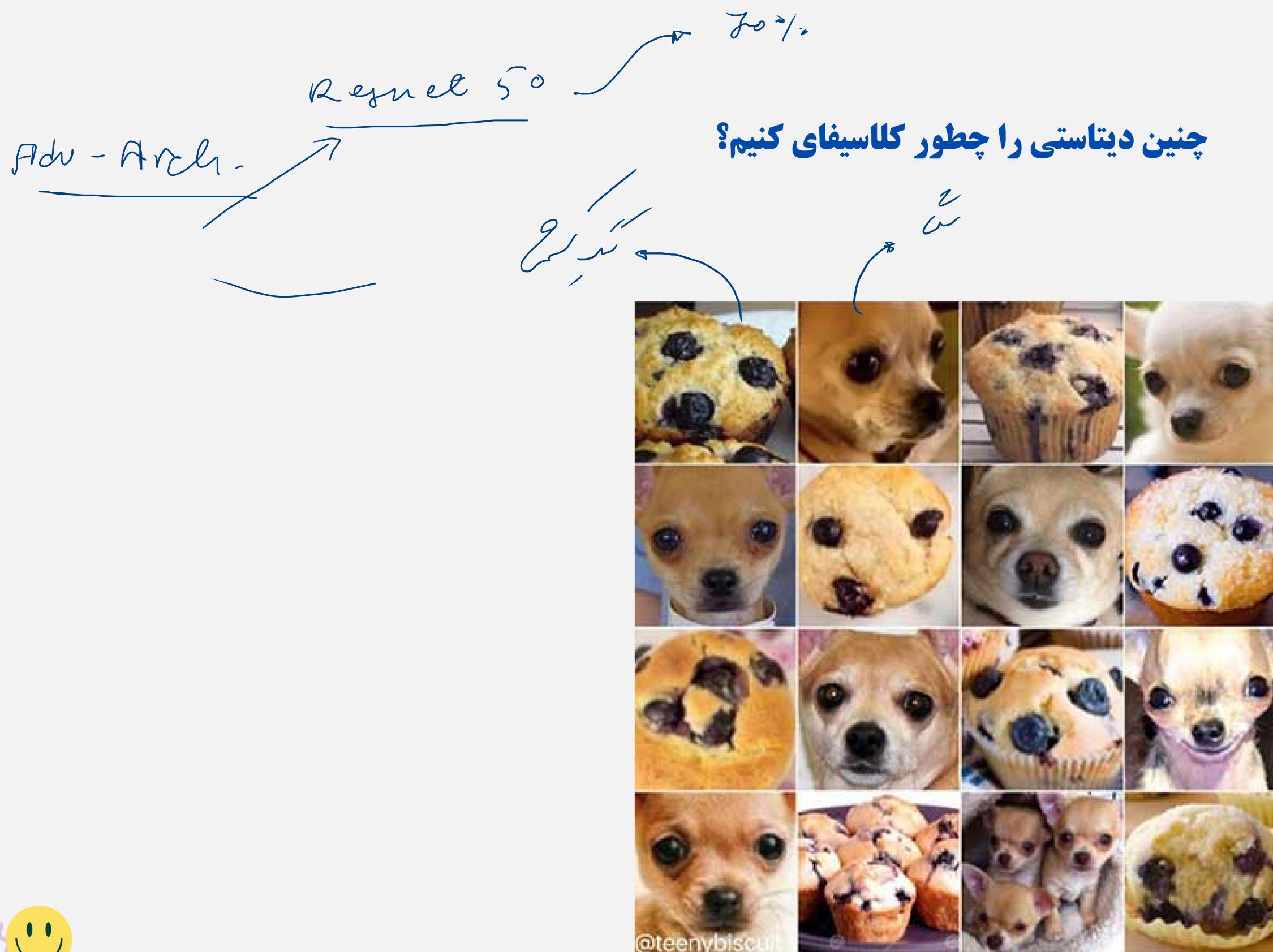


سوال

برای دسته بندی تصاویر بسیار مشابه باید چکار کنم؟

کلاسیفیکیشن (دسته بندی)



FUN

۲

IRIS

برای دسته بندی انواع گل زنبق (بسیار مشابه) باید چکار کنم؟

iris setosa



petal sepal

iris versicolor

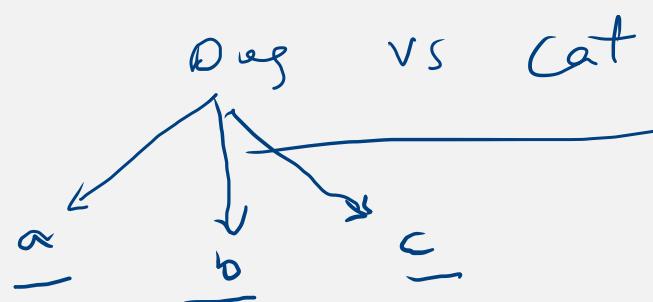


petal sepal

iris virginica



petal sepal



برای دسته بندی نژادهای سگ (بسیار مشابه) باید چکار کنم؟

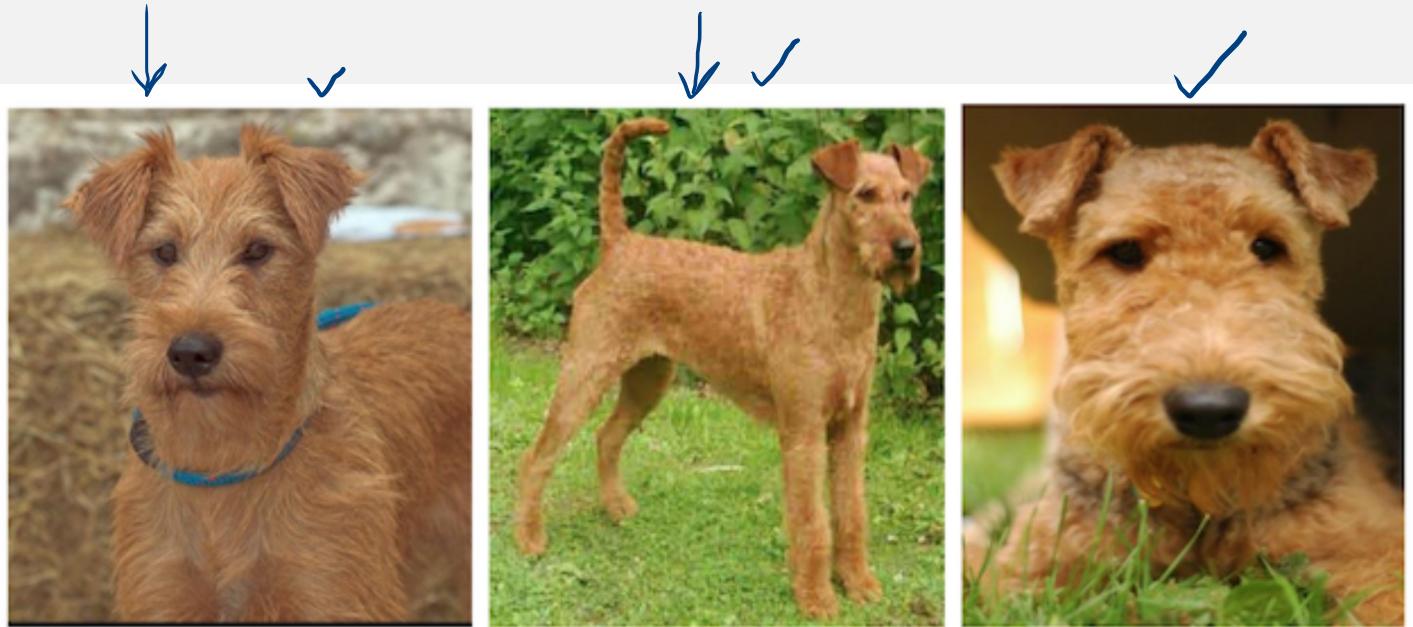


Fig 1(a): Irish Terrier

Fig 1(b): Fox Terrier

Fig 1(c): Airedale



Fig 2(a): Siberian Husky

Fig 2(b): Eskimo Dog

Fig 2(c): Malamute

برای دسته بندی تصاویر بسیار مشابه (نژادهای سگ، انواع گل زنبق، گونه های پرنده) باید چکار کنم؟

روش های پیشنهادی دانشجویان

swish Transformer

Rest of ResNet

PCA

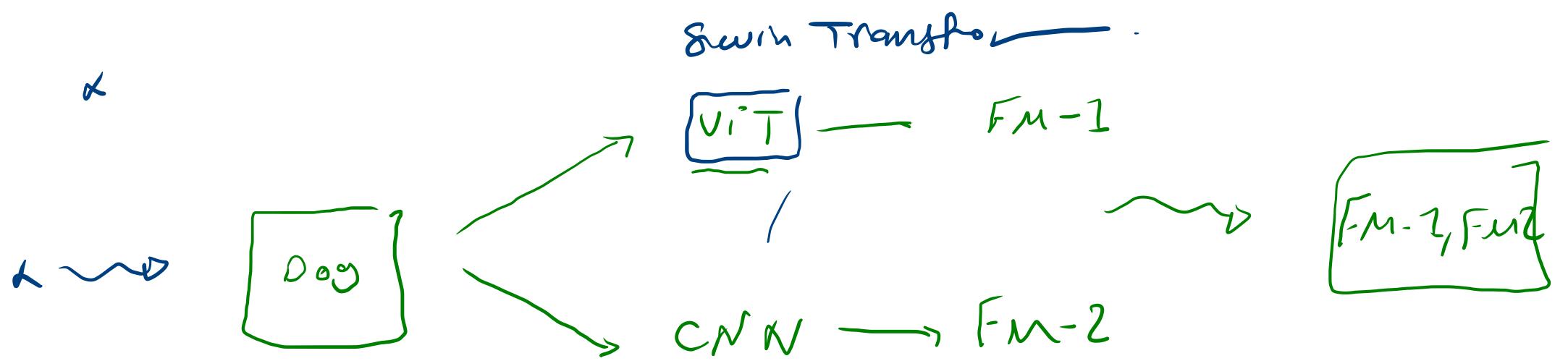
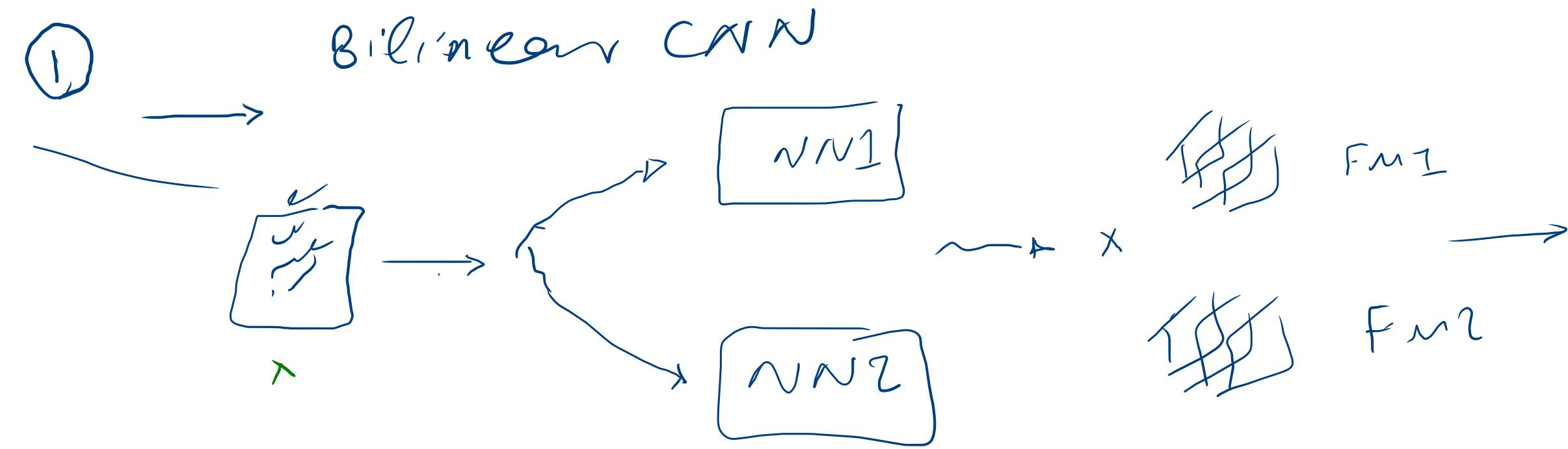
[Triplet loss, frame Met. \rightarrow عرضه \leftarrow تردد] • ✓
 کرنل برای سیستم \rightarrow \leftarrow مجموع \leftarrow adv-arch. • ✗

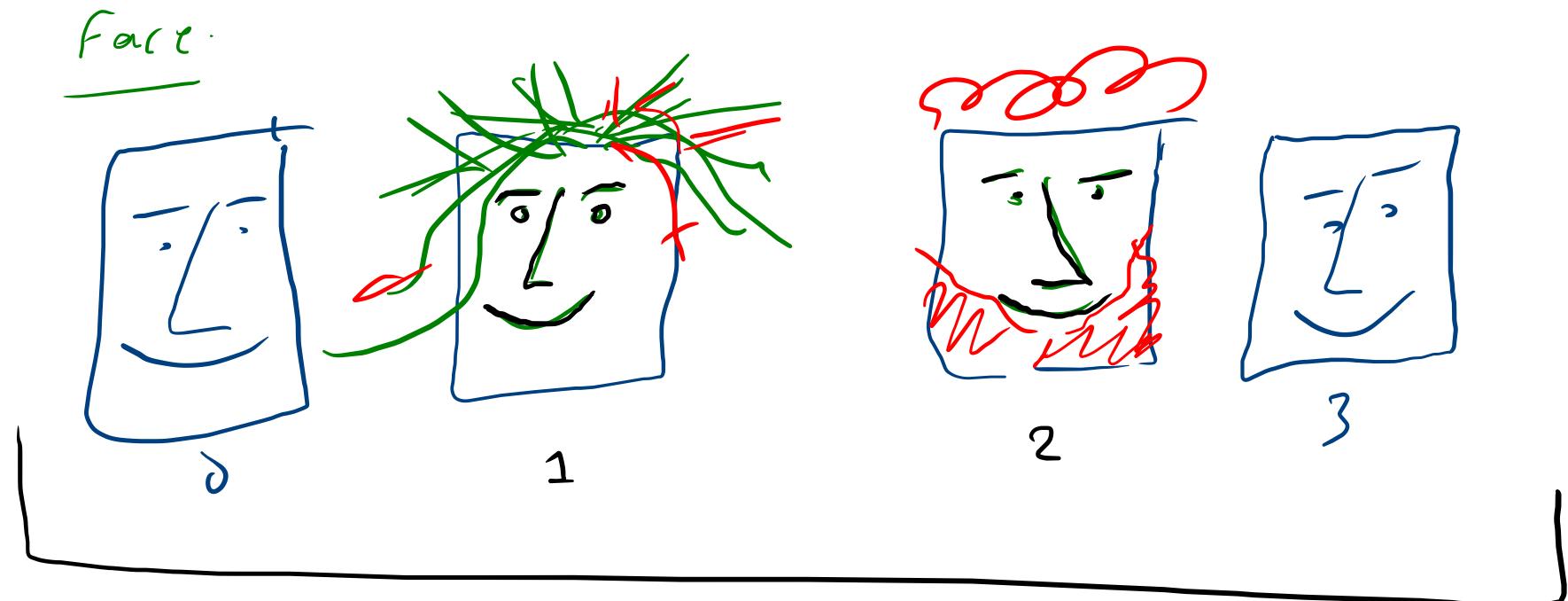
yolo \leftarrow سفرگرد \leftarrow تحریک \leftarrow باز تردد \leftarrow نت \leftarrow VIT • ✗

VIT \leftarrow ایندیا • ✓

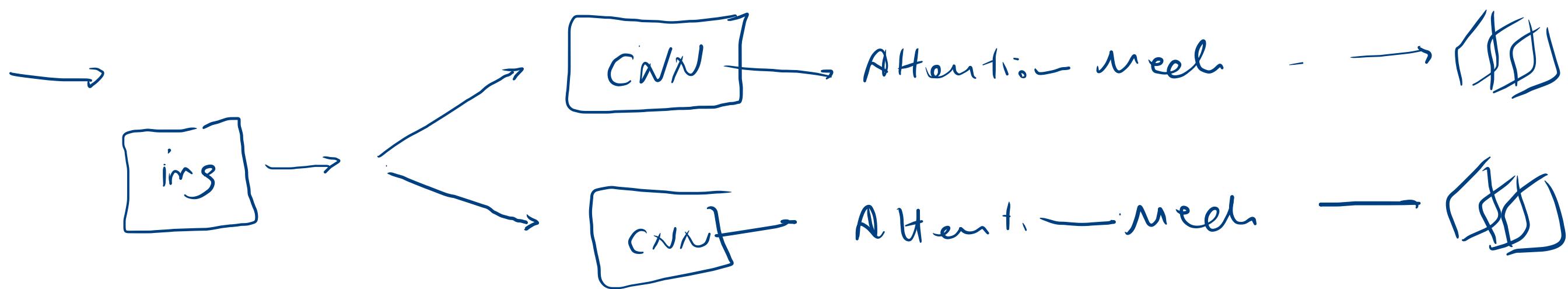
GCN \leftarrow سیستم \leftarrow نت •

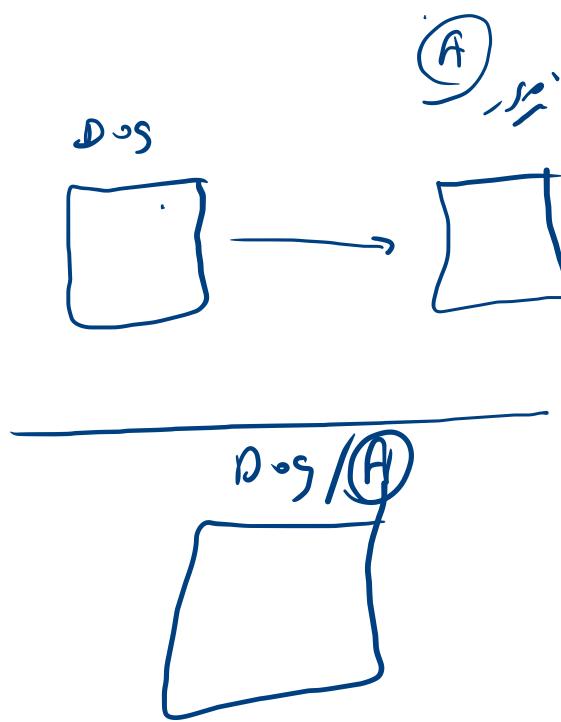
Attention mechanism \leftarrow نهادهای سگ \leftarrow نهادهای پرنده •





Bilinear CNN + Attention!





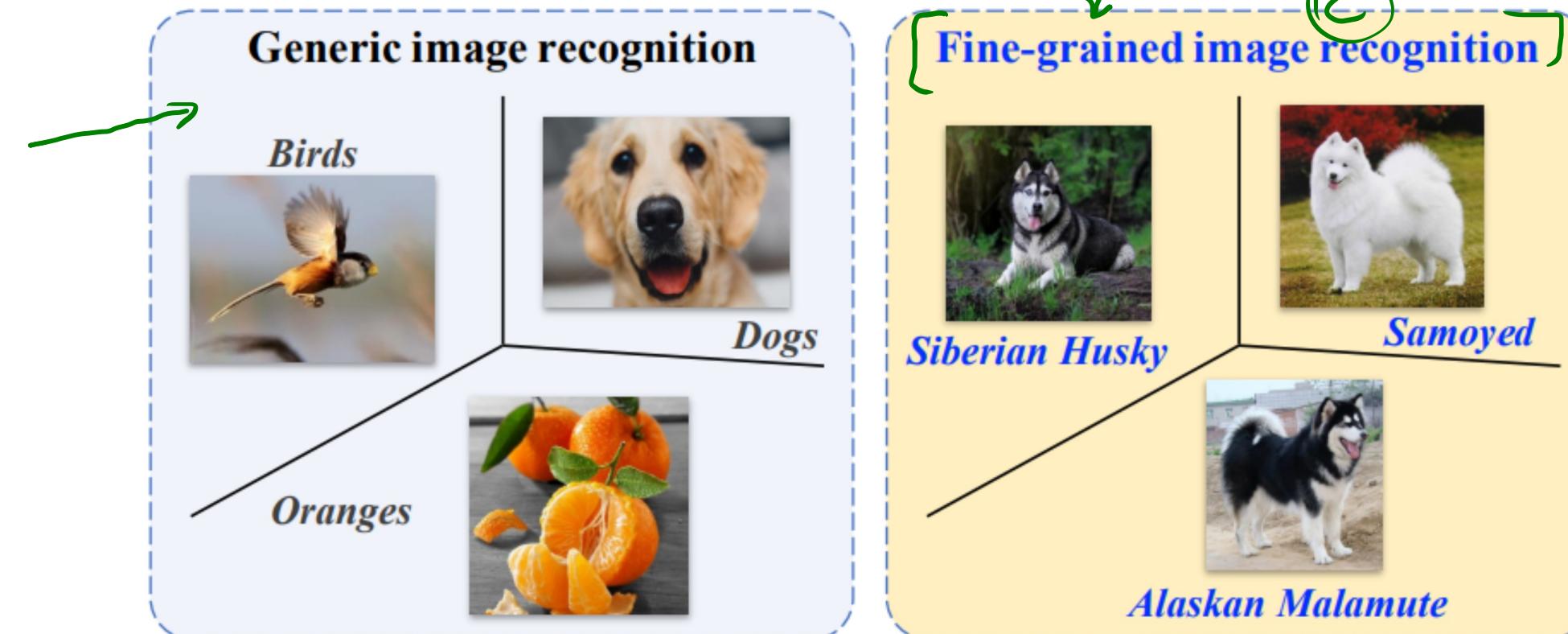
FGIC

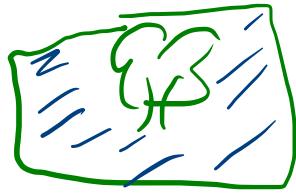
Fine Grained Image Classification

FGIC (Fine-Grained Image Classification)

شناصایی و دسته‌بندی تصاویر در سطوح جزئی‌تر و دقیق‌تر

- تشخیص تراشهای مختلف سگ (مانند گلدن رتریور در مقابل لابرادر رتریور)
- شناخت گونه‌های پرندگان (مثلاً گنجشک خانگی در مقابل گنجشک سفیدابرو)
- طبقه‌بندی مدل‌های خودرو (مثلاً BMW سری 3 در مقابل سری 5)
- تشخیص ~~نمودگان~~ نقوشی‌ها (مثلاً آثار داوینچی در مقابل رامبرانت)





چالش های FGIC



• تفاوت های جزئی: ویژگی های تفکیک کننده بسیار کوچک هستند (مثلاً تغییرات در رنگ پرهای پرندگان).

• عدم تعادل داده ها: معمولاً برخی کلاس ها داده های کمتری نسبت به بقیه دارند.

• پس زمینه های پیچیده: گاهی اشیا در شرایط نوری مختلف یا زوایای نامتعارف دیده می شوند.

مهمنه ترین چالش:

حتی یک بخش کوچک از شی می تواند تأثیر زیادی در طبقه بندی داشته باشد.



Artic Tern

Common Tern

Forster's Tern

Caspian Tern



Fig 1(a): Irish Terrier

Fig 1(b): Fox Terrier

Fig 1(c): Airedale

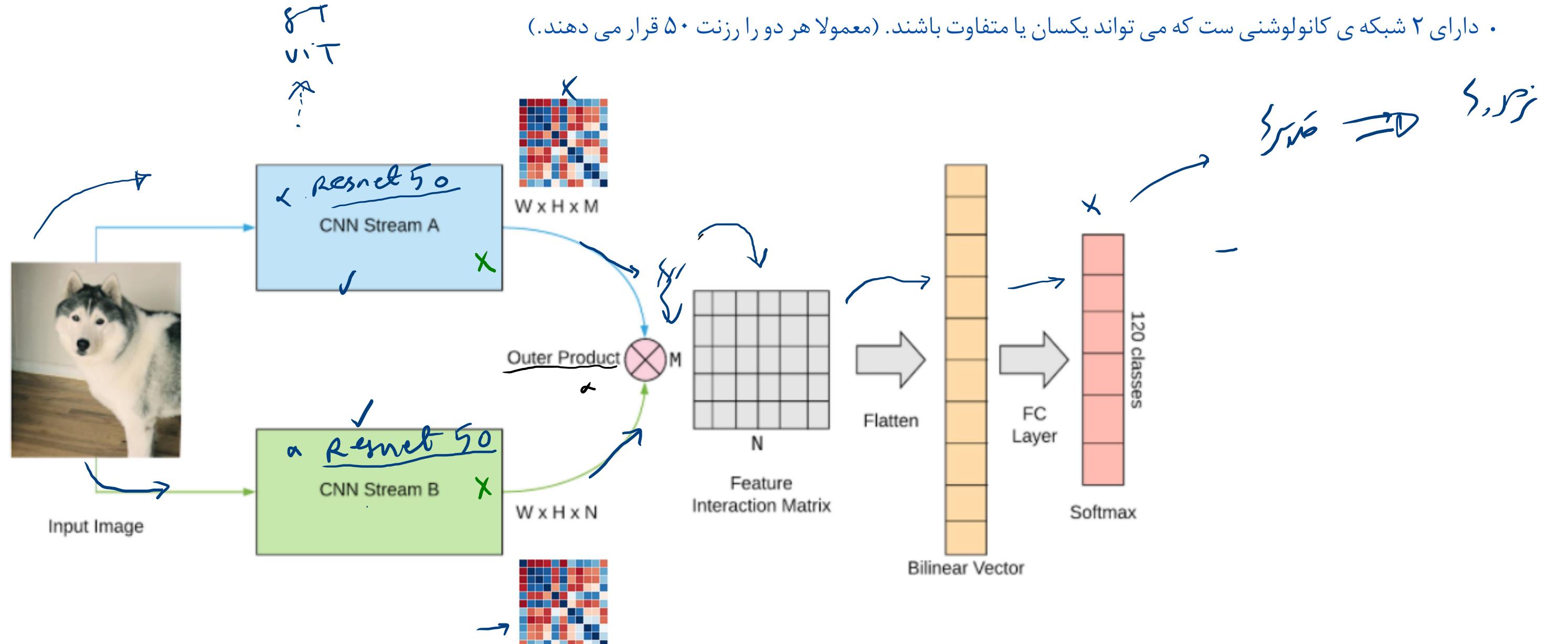


Fig 2(a): Siberian Husky

Fig 2(b): Eskimo Dog

Fig 2(c): Malamute

راه حل : [Bilinear CNN Model Architecture]

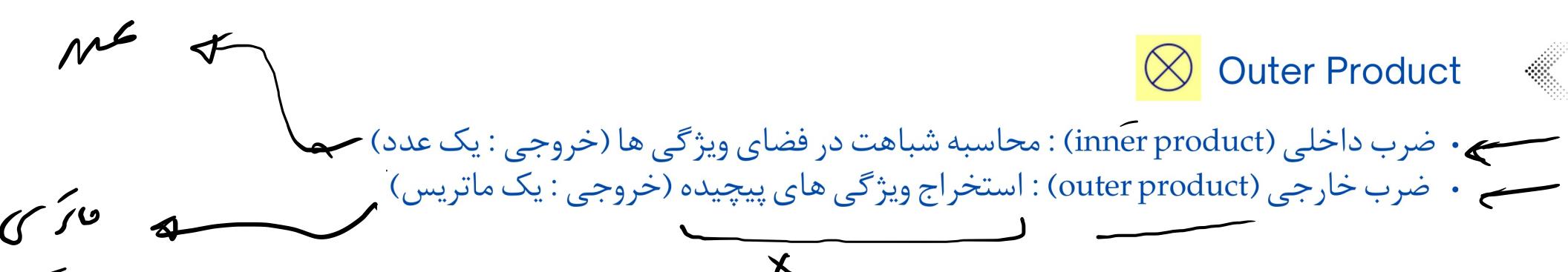


<https://arxiv.org/pdf/1504.07889.pdf>

<https://wandb.ai/authors/bcnn/reports/Fine-Grained-Image-Classification-FGIC-with-B-CNNs---Vmildzo0NDQ1Nzc>

$$a \cdot b = |a||b|\cos\theta$$

a

 Outer Product


$$\underline{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}, \quad \underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

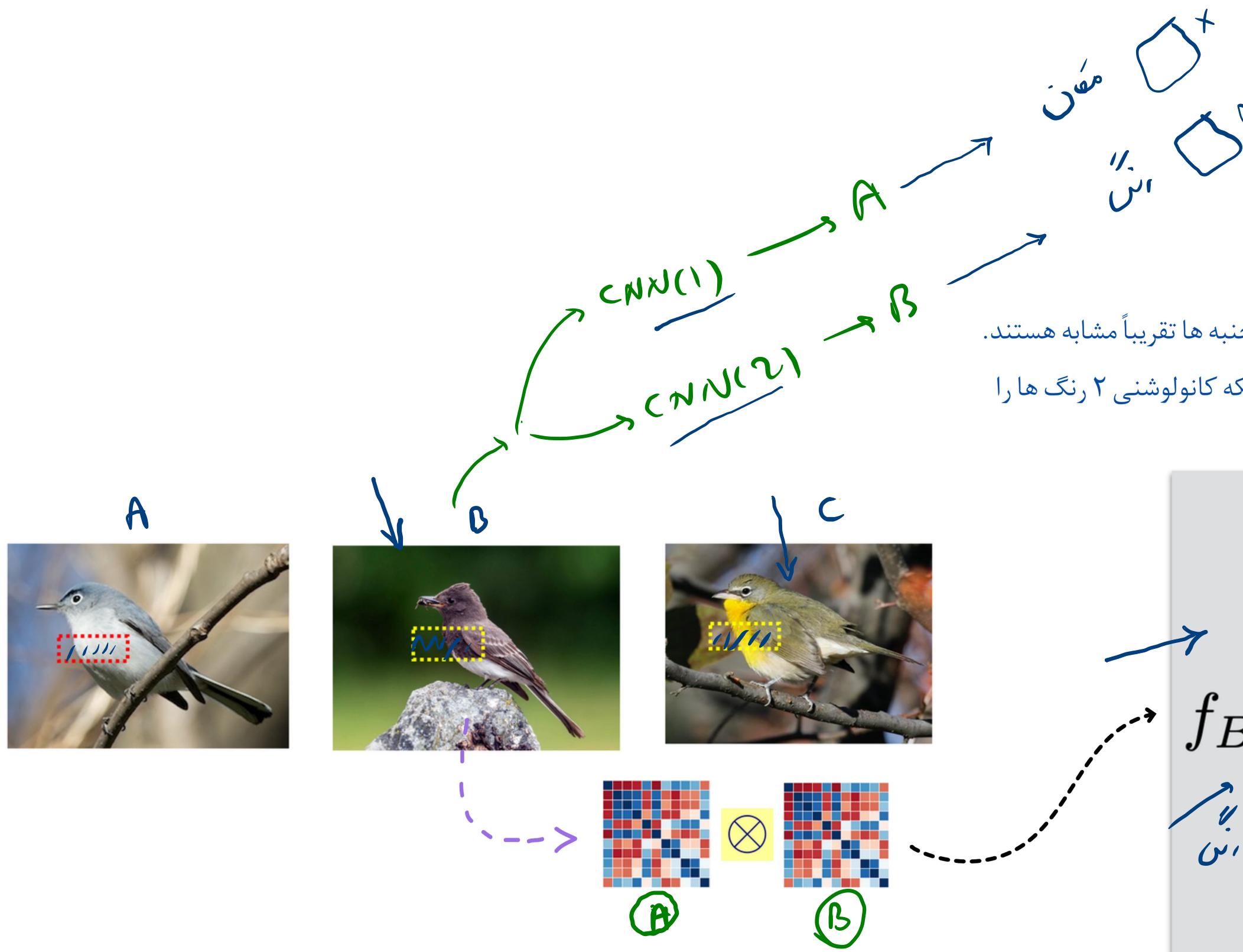
u v

~~Inner Product~~

$$\underline{u} \otimes \underline{v} = \underline{u}\underline{v}^T = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} [v_1 \quad v_2 \quad v_3] = \begin{bmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \\ u_3 v_1 & u_3 v_2 & u_3 v_3 \\ u_4 v_1 & u_4 v_2 & u_4 v_3 \end{bmatrix}.$$

$$\underline{u} \otimes \underline{v} = \mathbf{A} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \dots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \dots & u_2 v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_m v_1 & u_m v_2 & \dots & u_m v_n \end{bmatrix}$$

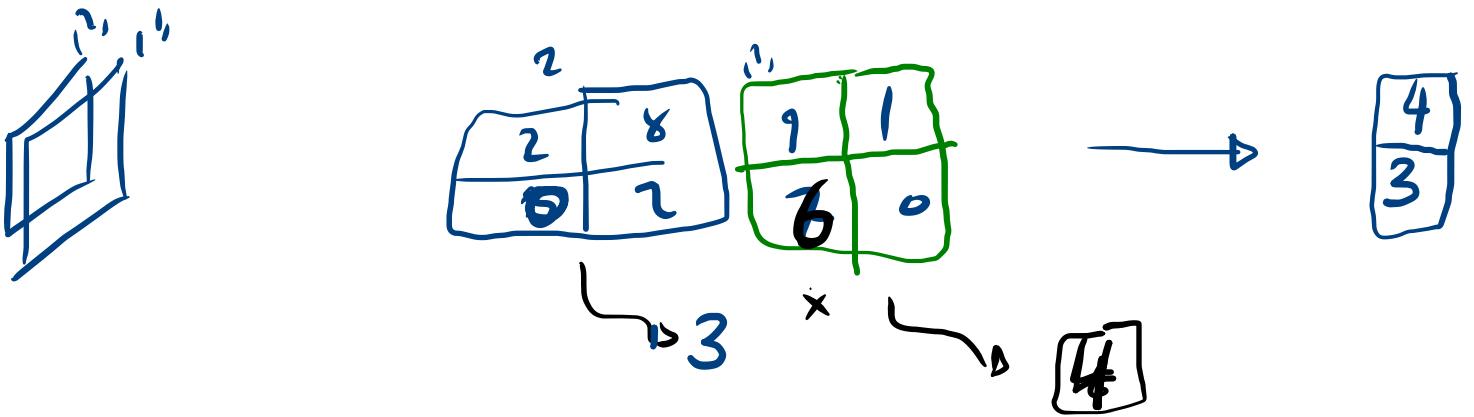
$$v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \rightarrow u \otimes v = \begin{bmatrix} 1 \\ 2 \end{bmatrix} [3, 4] = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$



BCNN Feature Interaction Matrix

	beak	tail	belly	legs	belly
Yellow red				/ / / /	
blue					
gray				— — — —	—
blue					
black					

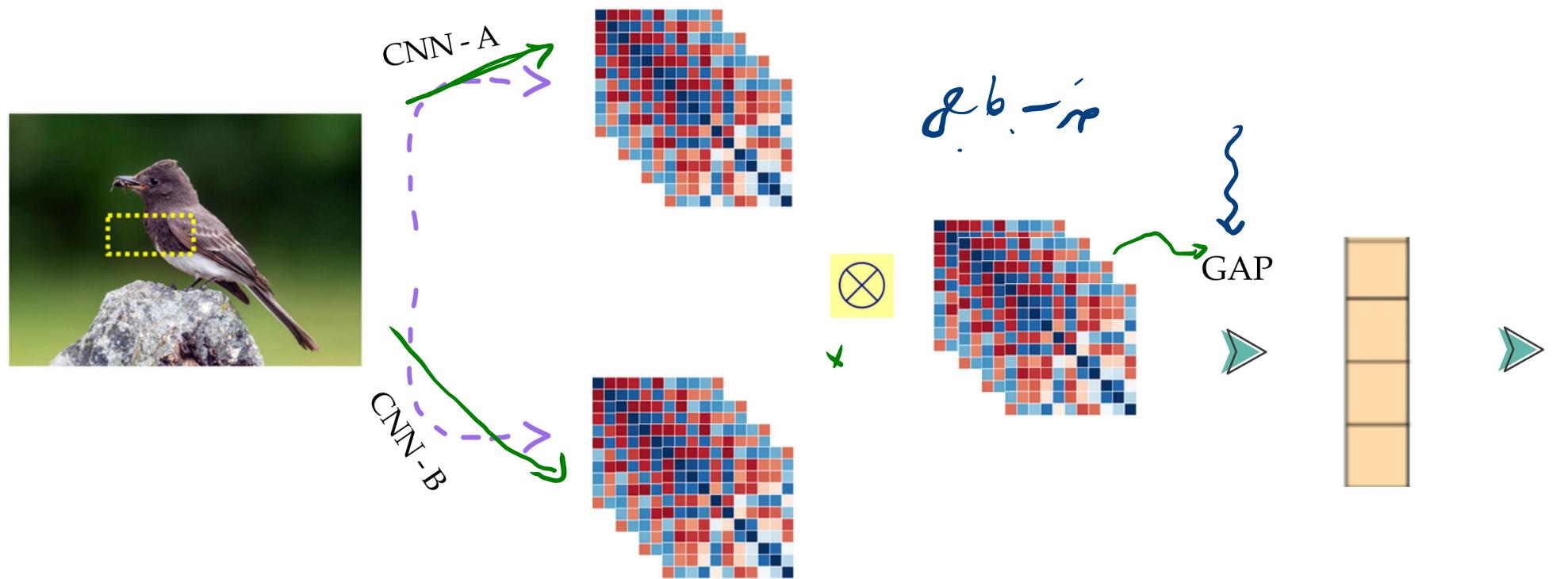
example “gray belly”



Bilinear Pooling

گام اول : ضرب خارجی

گام دوم : Global Average Pooling



FGIC : کلاسیفیکیشن تصاویر مشابه

$$x = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \rightarrow y = ? \rightarrow \text{Sign}(x) = \begin{bmatrix} \text{Sign}(4) \\ \text{Sign}(3) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow q$$

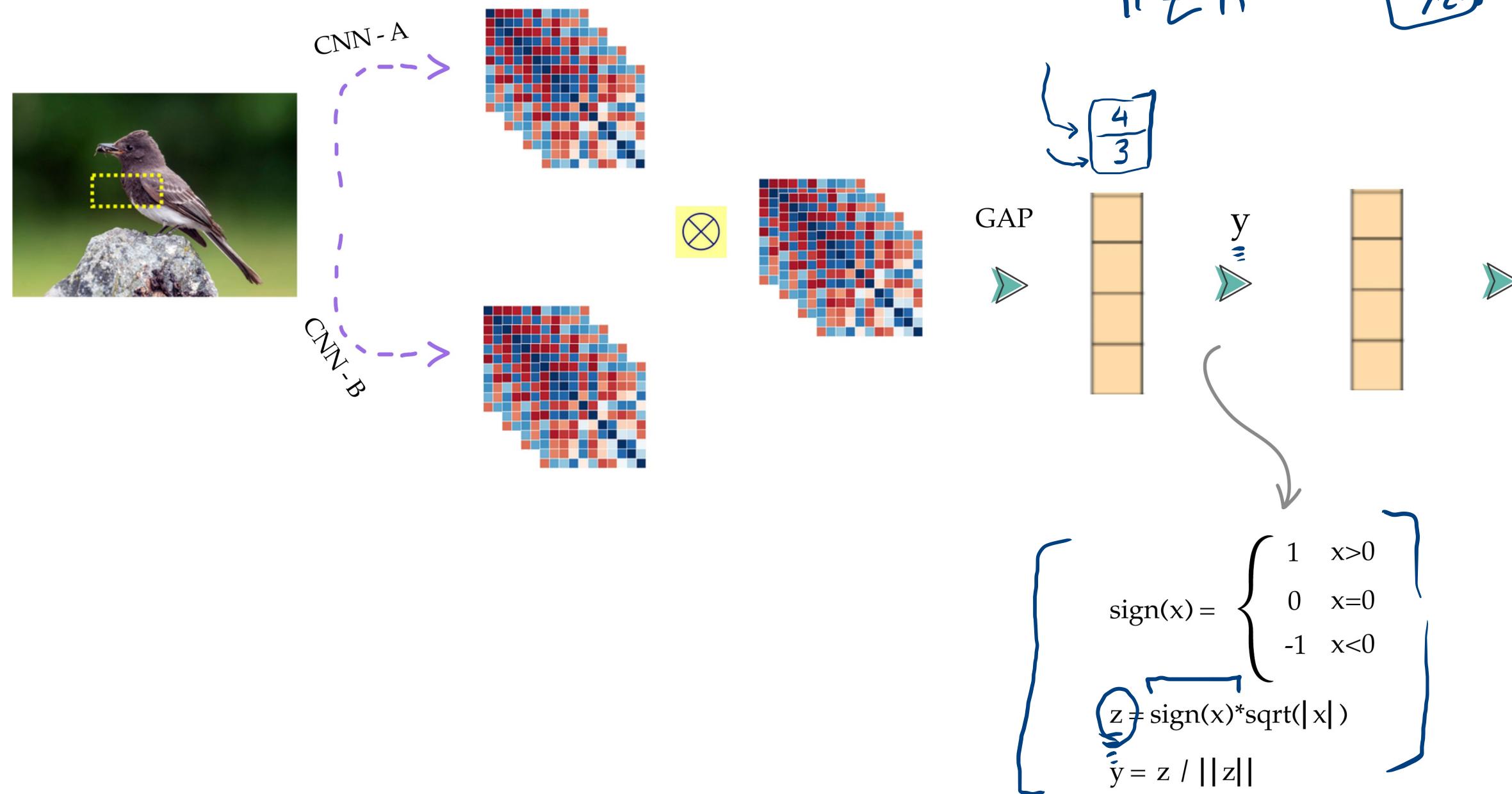
$$z = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} \sqrt{14} \\ \sqrt{13} \end{bmatrix} = \begin{bmatrix} 2 \\ 14 \end{bmatrix} \rightsquigarrow z = \begin{bmatrix} 2 \\ 14 \end{bmatrix} \quad \|z\| = \sqrt{4+3} = \sqrt{7} = 2.5$$

$$\text{Bilinear Pooling}$$

گام اول: ضرب خارجی

گام دوم: Global Average Pooling

گام سوم: L2 normalization



لایه FC

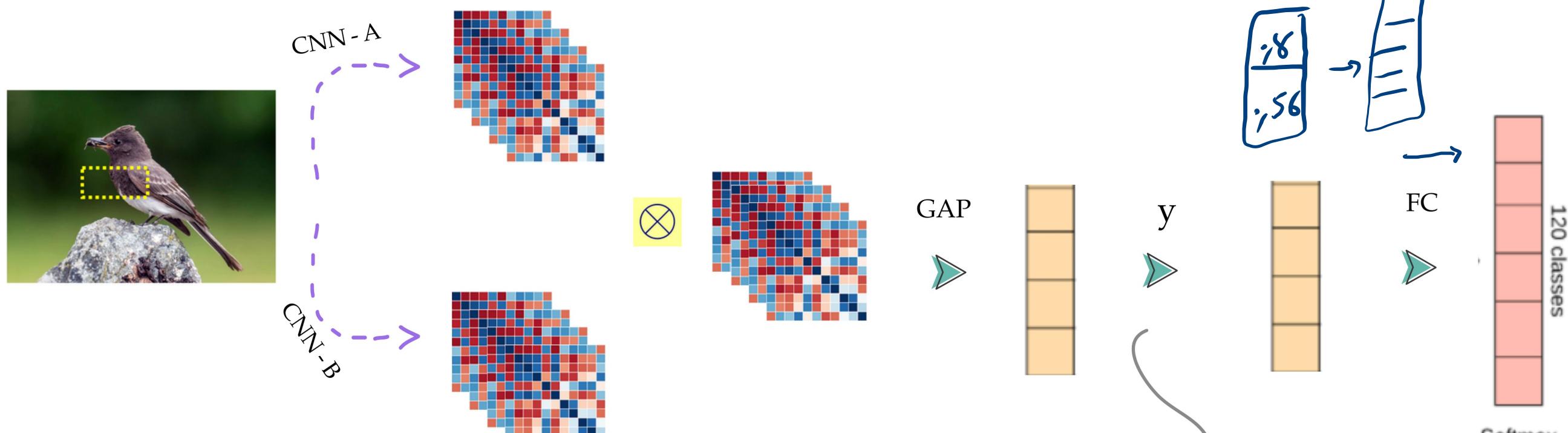
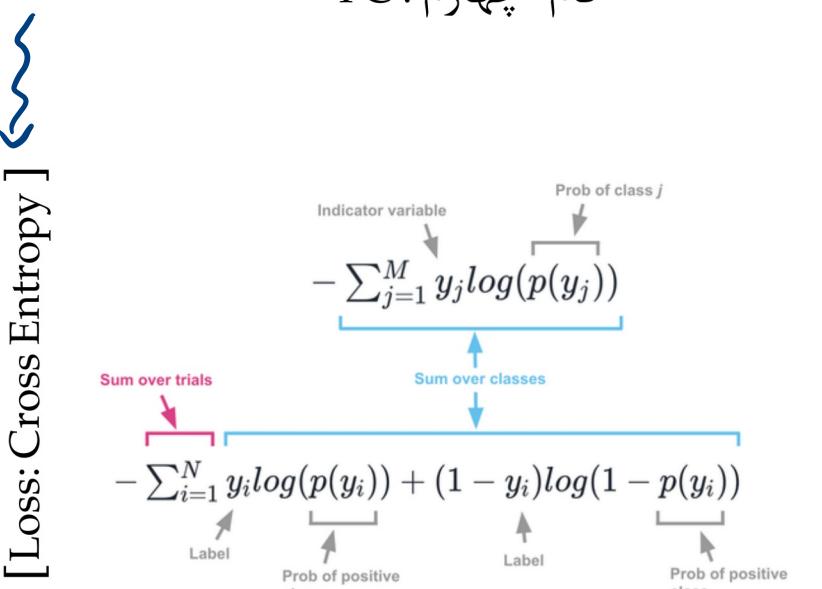


گام اول : ضرب خارجی

گام دوم : Global Average Pooling

گام سوم : L2 normalization

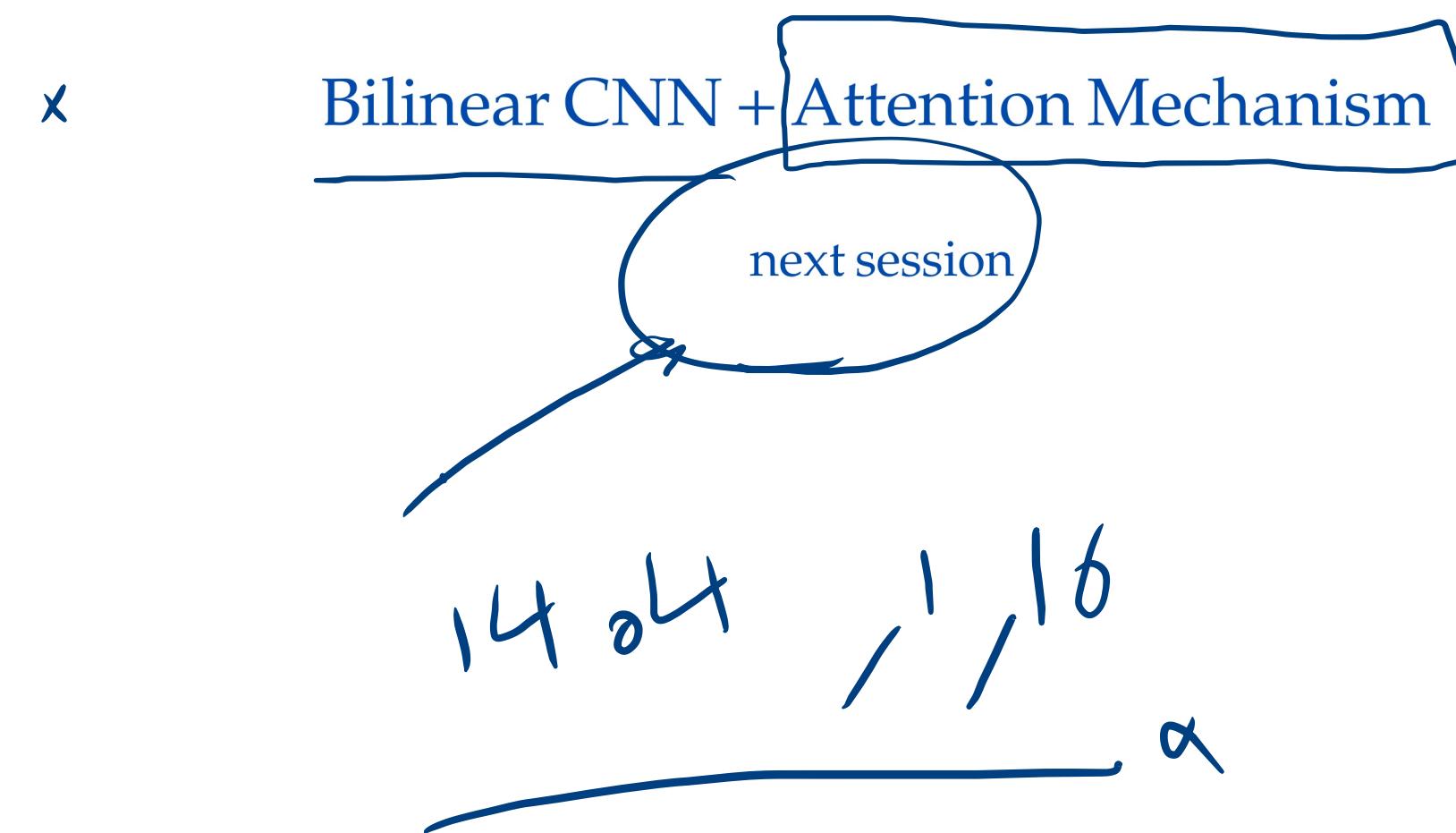
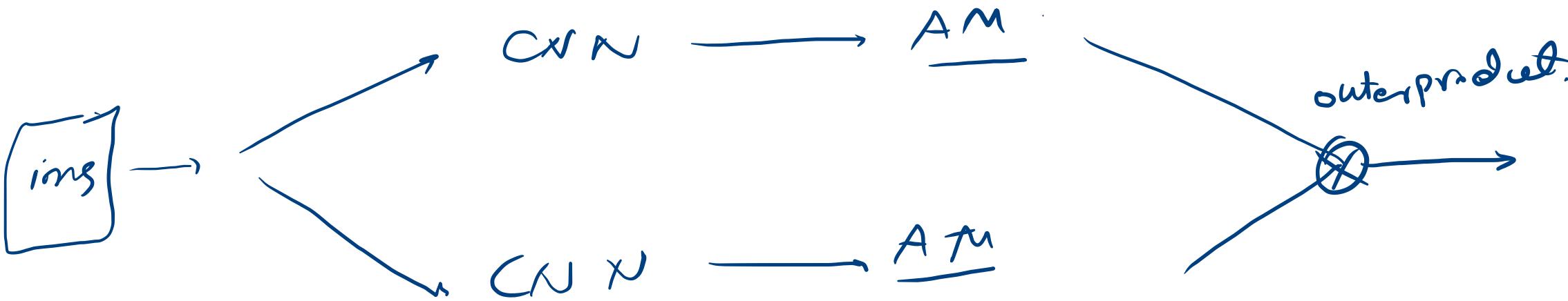
گام چهارم : FC



$$\text{sign}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

$$z = \text{sign}(x) * \sqrt{|x|}$$

$$y = z / \|z\|$$



The End

With Love