

کارگاه یادگیری عمیق با پایتون: بازشناسی چهره

سید ناصر رضوی www.snrazavi.ir

۱۳۹۷

- انواع وظایف مرتبط با تصاویر:
 - دسته‌بندی، مکان‌یابی، شناسایی اشیا و قطعه‌بندی
 - دسته‌بندی تصویر
 - مسئله دسته‌بندی سگ و گربه
 - مجموعه داده
 - تعریف و ایجاد مجموعه داده‌ها در پای‌تورچ
 - انجام انواع تبدیلات بر روی تصاویر و داده‌افزایی
 - تعریف و ایجاد بارگذار داده‌ها
- پیاده‌سازی و آموزش یک شبکه کانولوشنی ساده
- یادگیری انتقالی
- استفاده از شبکه‌هایی که قبلاً برای انجام یک وظیفه مرتبط، آموزش یافته‌اند

یادآوری: تمرین اول

۳

□ دسته‌بندی نهال ۱۲ نوع گیاه.



□ بهترین نتایج.

□ حامد همایی راد: ۹۸/۳۶

□ محمد زینالی: ۹۷/۹۸

□ رقیه یوسفی: ۹۶/۲۲

یادآوری: اصلاح کد

۴

```
def train_model(model, train_dl, valid_dl, criterion, optimizer,
                scheduler=None, num_epochs=10):

    if not os.path.exists('models'):
        os.mkdir('models')

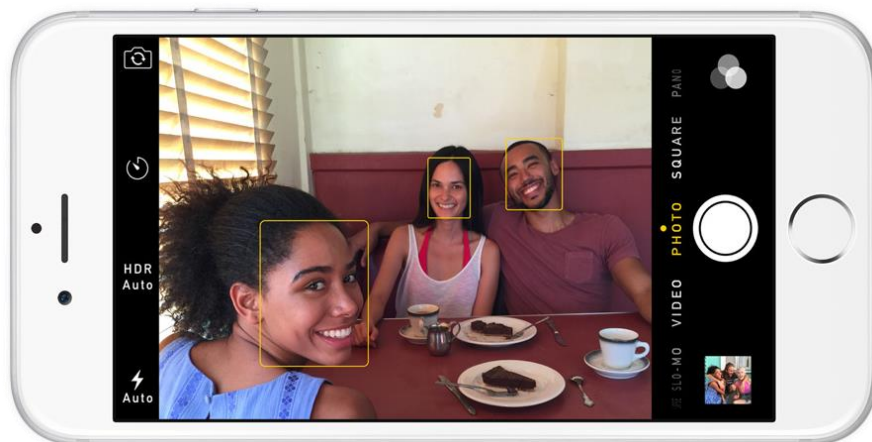
    since = time.time()

    best_model_wts = model.state_dict()
    best_acc = 0.0

    for epoch in range(num_epochs):
        print('Epoch {}/{}'.format(epoch + 1, num_epochs))
        print('-' * 10)

        ## train and validate
        model = train_one_epoch(model, train_dl, criterion, optimizer, scheduler)
        val_acc = validate_model(model, valid_dl, criterion)

        # deep copy the model
        if val_acc > best_acc:
            best_acc = val_acc
            best_model_wts = model.state_dict().copy()
            torch.save(best_model_wts, "./models/epoch-{}-acc-{:0.5f}.pth".format(epoch, best_acc))
```



□ دسته‌بندی تصاویر

□ مسئله بازشناسی چهره

■ تعداد دسته‌ها نسبتاً زیاد

■ تعداد نمونه‌های آموزشی مربوط به هر دسته بسیار کم

□ شبکه‌های سیامی

□ تابع هزینه سه‌تایی

□ تابع هزینه دودویی

■ حل مسئله به عنوان یک مسئله دسته‌بندی دودویی

دسته‌بندی تصاویر: یادگیری تک شات

۶

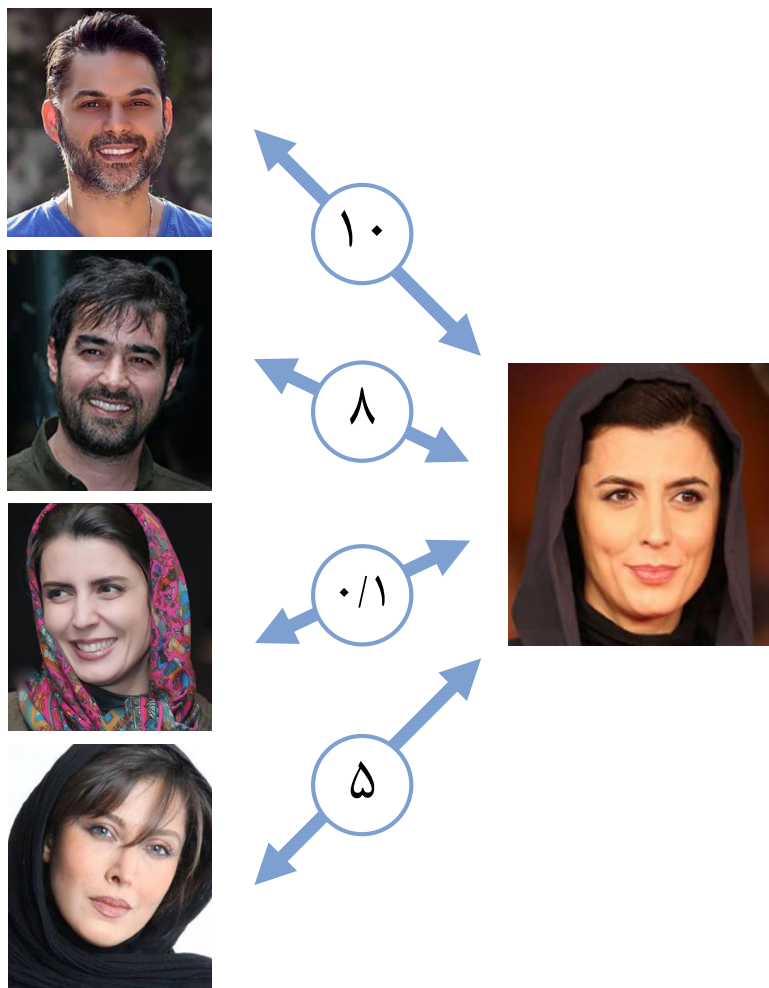
□ یادگیری از روی یک نمونه آموزشی [یا یک تعداد بسیار کم از نمونه‌های آموزشی].



یادگیری معیار شباهت

۷

□ تابع فاصله.



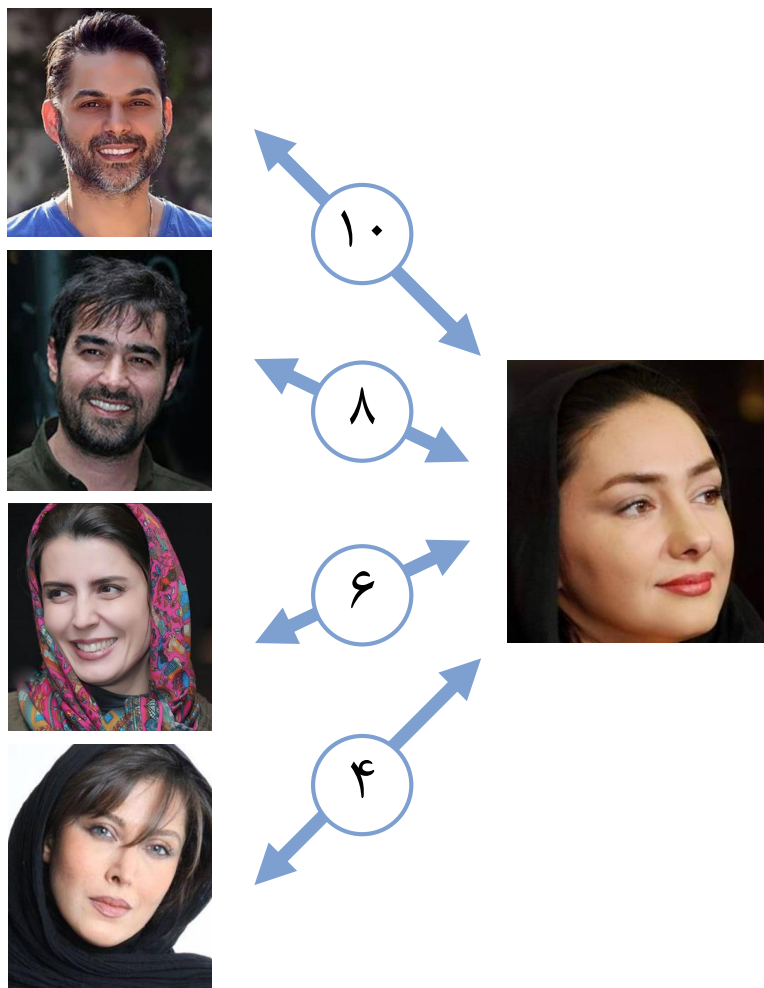
$$d(\text{Image 1}, \text{Image 2}) \leq \tau$$

$$d(\text{Image 3}, \text{Image 4}) > \tau$$

یادگیری معیار شباهت

۸

□ تابع فاصله.



$$d(\text{Image 1}, \text{Image 2}) \leq \tau$$

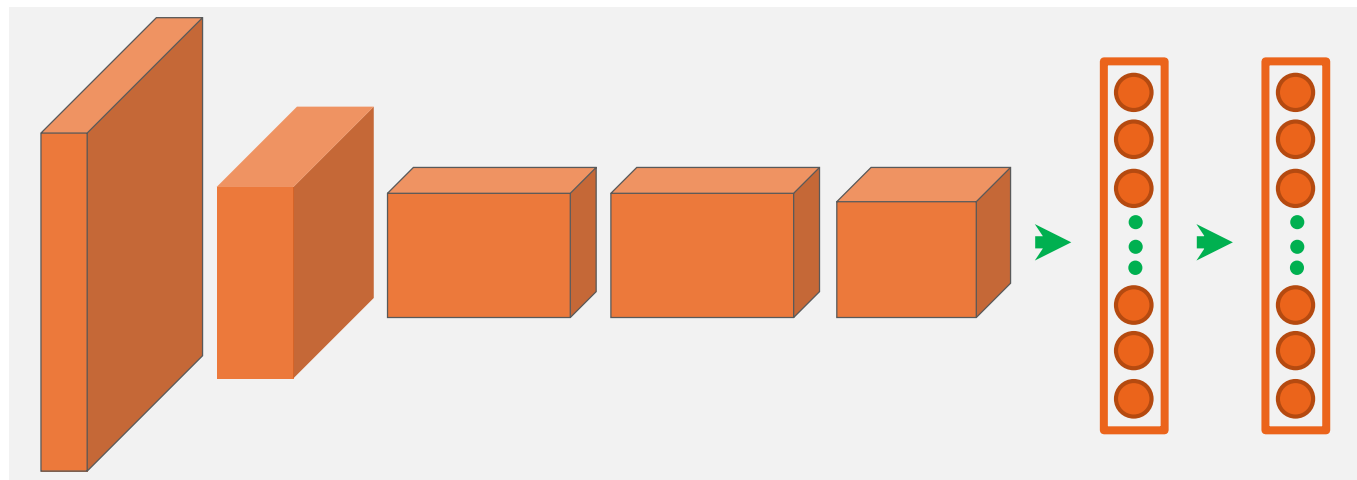
$$d(\text{Image 1}, \text{Image 3}) > \tau$$

شبکه سیامی

۹



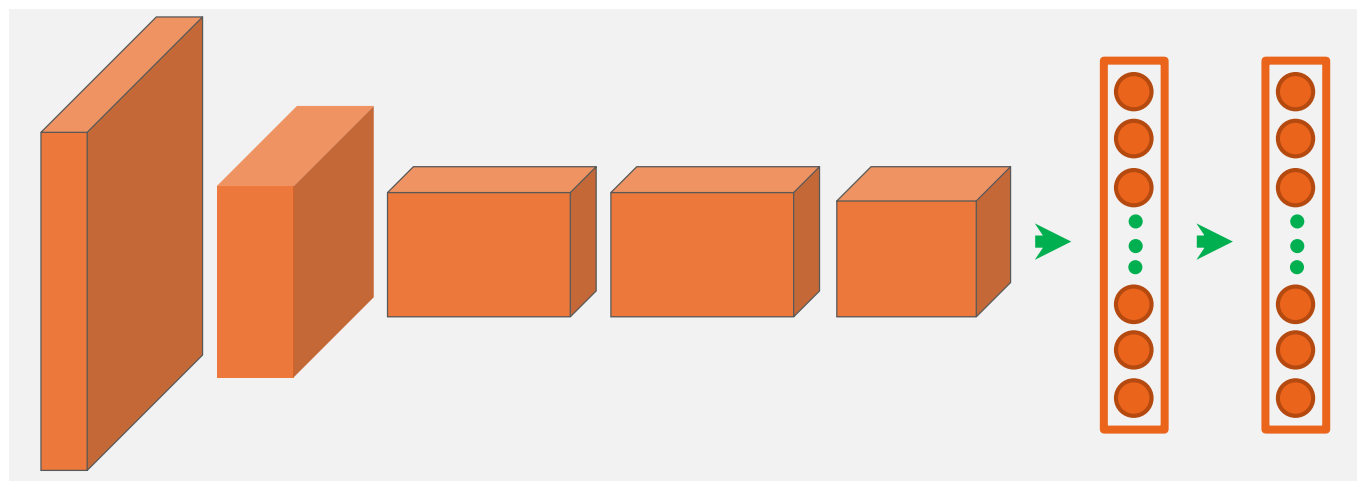
$x^{(1)}$



بدرار ویژگی [۱۲۸ بُعد]
 $\rightarrow f(x^{(1)})$



$x^{(2)}$



$\rightarrow f(x^{(2)})$

شبکه سیامی: تابع فاصله

۱۰

□ محاسبه فاصله دو تصویر $x^{(i)}$ و $x^{(j)}$

$$d(x^{(i)}, x^{(j)}) = \|f(x^{(i)}) - f(x^{(j)})\|_2^2$$

□ آموزش پارامترهای شبکه کانولوشنی.

اگر دو تصویر $x^{(i)}$ و $x^{(j)}$ متعلق به یک شخص باشند:

$$\|f(x^{(i)}) - f(x^{(j)})\|^2 \rightarrow 0$$

اگر دو تصویر $x^{(i)}$ و $x^{(j)}$ متعلق به یک شخص نباشند:

$$\|f(x^{(i)}) - f(x^{(j)})\|^2 \rightarrow \infty$$

شبکه سیامی: آموزش

۱۱

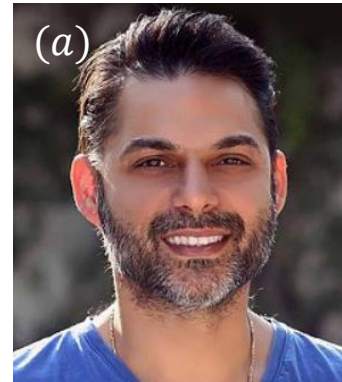
□ تابع هزینه سه گانه.



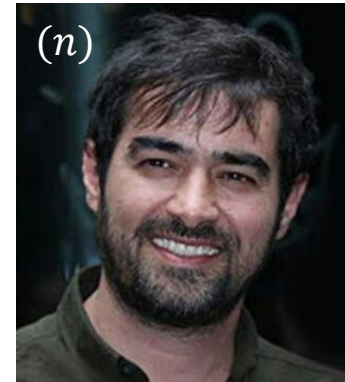
لنگر



مثبت



لنگر



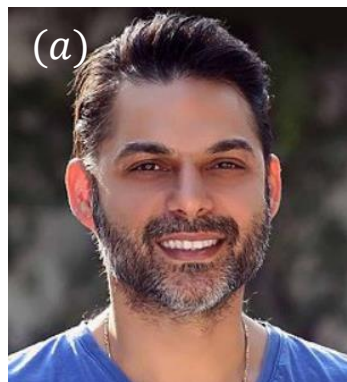
منفی

$$\|f(a) - f(p)\|^2 \leq \|f(a) - f(n)\|^2$$

شبکه سیامی: آموزش

۱۲

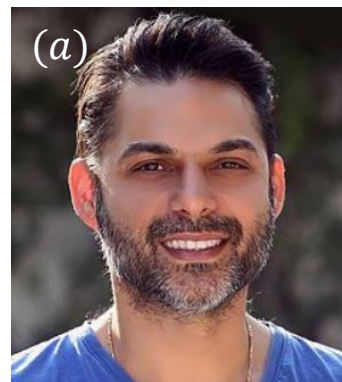
□ تابع هزینه سه گانه.



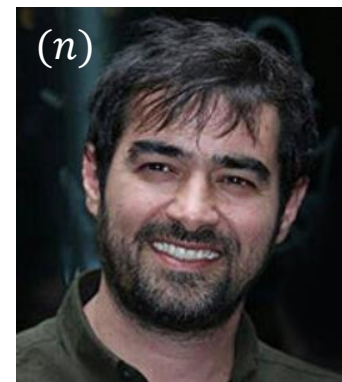
لنگر



مثبت



لنگر



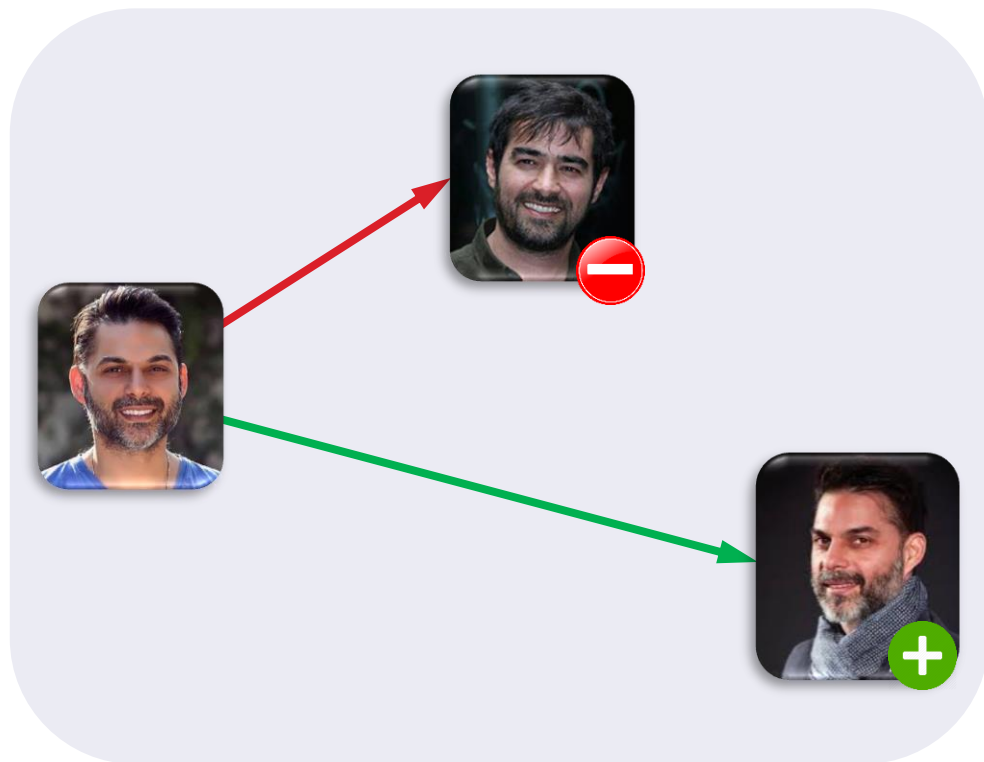
منفی

$$\|f(a) - f(p)\|^2 - \|f(a) - f(n)\|^2 + \alpha \leq 0$$

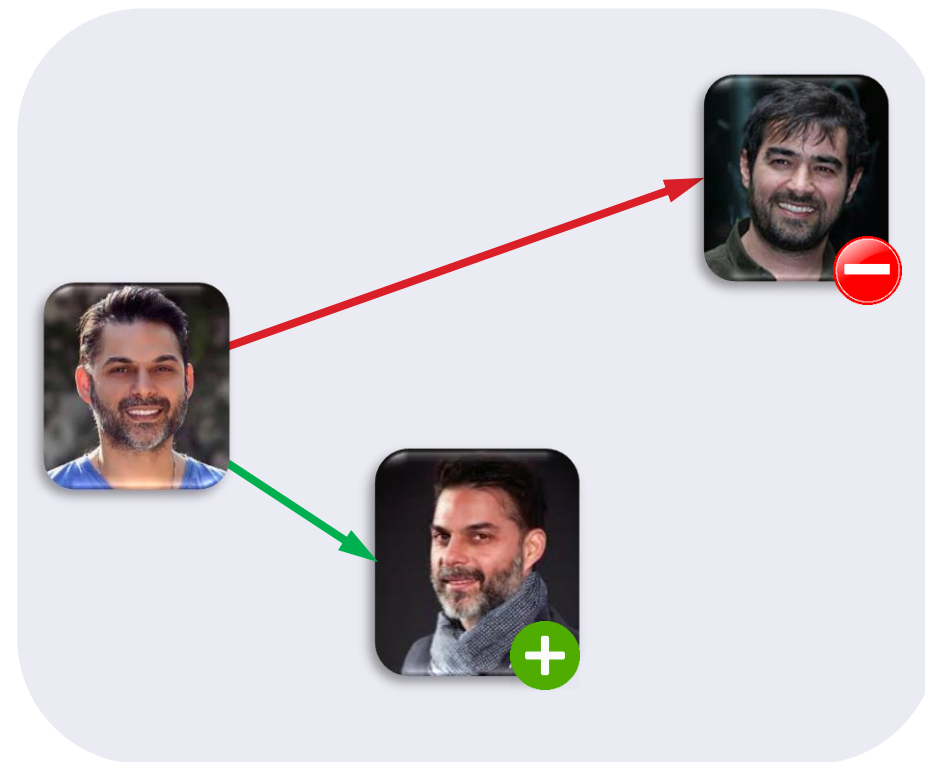
اجتناب از پاسخ های بدیهی!

شبکه سیامی: آموزش

۱۳



یادگیری



یک نمونه آموزشی (شامل تصاویر لنگر، مثبت و منفی)

شبکه سیامی: آموزش

۱۴

□ تابع هزینه سه گانه.

$$L(a, p, n) = \max(\|f(a) - f(p)\|^2 - \|f(a) - f(n)\|^2 + \alpha, 0)$$

□ انتخاب یک مجموعه از سه تایی های مناسب برای مجموعه آموزشی.

□ انتخاب سه تایی هایی که یادگیری آنها برای شبکه عصبی **دشوار** است!

$$d(a, p) \cong d(a, n)$$

شبکه سیامی: مجموعه آموزشی

۱۵

لنگر



مثبت



منفی



⋮



⋮



⋮

شبکه سیامی: نتایج

۱۶



شبکه سیامی: آموزش

۱۷



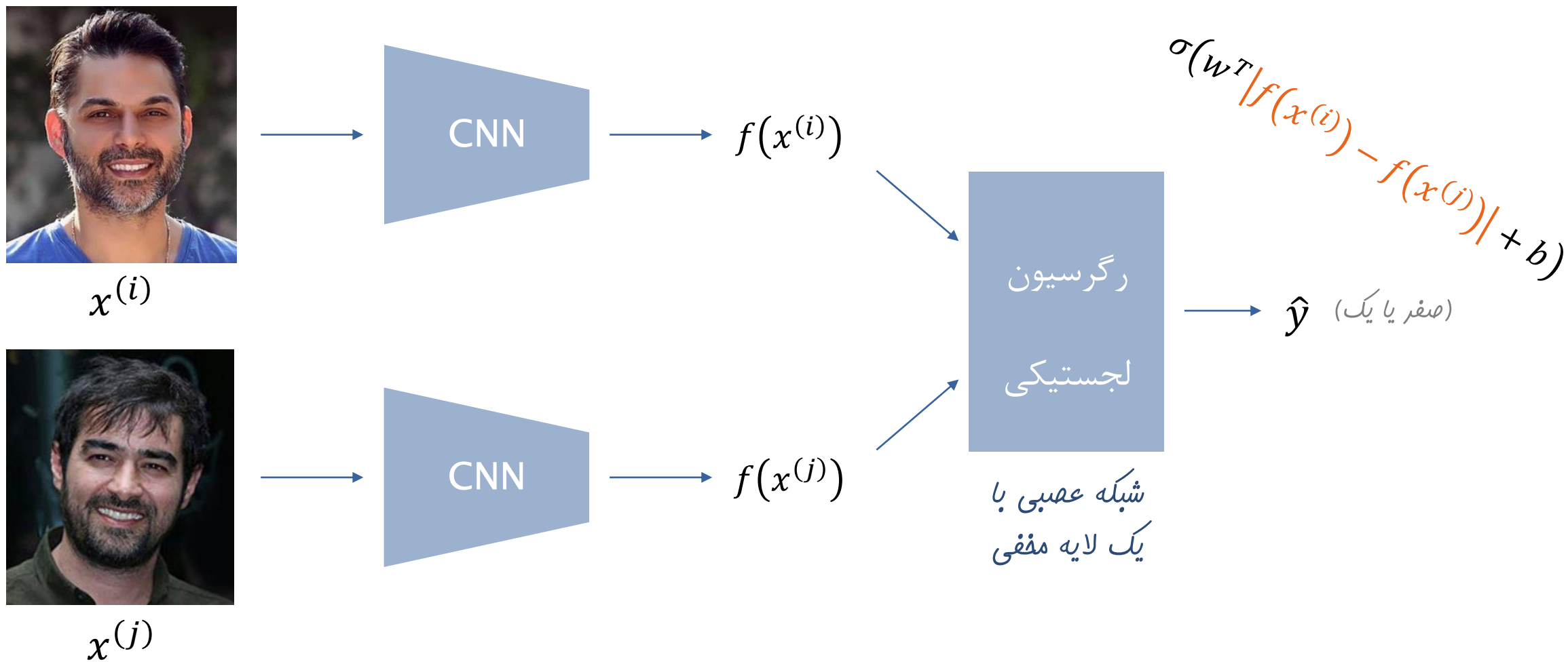
شبکه سیامی: آموزش

۱۸



تابع هزینه دودویی

۱۹



تابع هزینه دودویی: مجموعه داده

۲۰



|

|

|

•

|

•

•

•

چالش شناسایی نهنگ (کگل)

۲۱



□ تعداد ۹۸۵۰ تصویر

□ بیش از ۳۰۰۰ نوع
نهنگ

