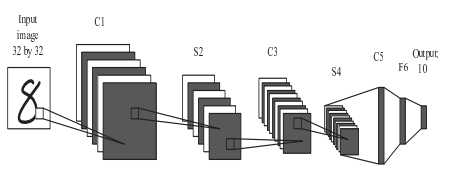
Background of gender classification:

Gender classification from face images has long been heated concern in biometric research. Traditional approaches are relied on hand-engineered features which can be grouped to be either global or local. The main issue with handcrafted features based approaches is that they require domain knowledge and may not generalize well. Because of that, we focus on approaches that utilize features learned from neural networks. Artificial feedforward neural networks have been around for decades for use in classification tasks,. In the 1990s, they began to be employed for gender classification[1]. However, the shallow structure of early neural networks has restrained their performance and applicability[2]. It was not until late 2012 when Krizhevsky *et al.* won the ImageNet Recognition Challenge with a ConvNet that neural networks gained attention again. In the following years, various deep nets were successfully applied to a variety of visual recognition tasks including facial gender classification. Verma *et* *al.* displayed that the CNN filters correspond to similarfeatures that neuroscientists identified as cues used by human beings to recognize gender[3]. Motivated by the dropout technique in training deep nets, Eidinger *et al*. trained a SVM with random dropout of some features and achieved promising results on their relatively small Adience dataset, on which Levi and Hassner later trained and tested a not-very-deep CNN. Instead of training on entire images, Mansanet *et al*. trained relatively shallow nets using local patches and reported better accuracy than whole image based nets of similar depths[4]. The proposal is mianly involved with CNN such as lenet5, Alexnet and VGG.

Content:Lenet5

Lenet5 is a common model of CNNs[5]. The input is the 3232 pixels image. The LeNet5 is consist of 7 layers: three convolutional layers, two subsampling layers, one fully connected layer and the output layer. The first layer is a convolutional layer (C1), which has 6 feature planes of 2828 pixels. In the subsampling layer(S2), these planes are reduced into 1414 pixels for one feature plane. The next convolutional layer (C3) extends the number of feature maps to sixteen. The subsampling layer S4 acts as S2, 16 feature planes are reduced to half their sizes. The last convolutional layer C5 has 120 feature planes, C5 is equal to a fully connected layer. The fully connected layer F6 contains 84 units connected to the 120 units of C5. Finally, the output layer is a Euclidean RBF layer of 10units. Lenet5 is shown in Fig.1.



Reference:

Fig.1 structure of Lenet5

[1]Q. Tian, T. Arbel and J. J. Clark, "Deep LDA-Pruned Nets for Efficient Facial Gender Classification," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, 2017, pp. 512-521.

[2]N. Yu, P. Jiao and Y. Zheng, "Handwritten digits recognition base on improved LeNet5," The 27th Chinese Control and Decision Conference (2015 CCDC), Qingdao, 2015, pp. 4871-4875.

[3]L. Cai, H. Zeng, J. Zhu, J. Cao, J. Hou and C. Cai, "Multi-view joint learning network for pedestrian gender classification," 2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), Xiamen, China, 2017, pp. 23-27.

[4]F. Becerra-Riera, H. Méndez-Vázquez, A. Morales-Gonzalez and M. Tistarelli, "Age and gender classification using local appearance descriptors from facial components," 2017 IEEE International Joint Conference on Biometrics (IJCB), Denver, CO, USA, 2017, pp. 799-804.

[5]T. Isbister, L. Kaati and K. Cohen, "Gender Classification with Data Independent Features in Multiple Languages," 2017 European Intelligence and Security Informatics Conference (EISIC), Athens, 2017, pp. 54-60.