1. (a) Data Augmentation for dealing with not all possible combinations are available.

⇒ Data Augmentation can create samples for the missing combinations during preprocessing, which helps the training model to recognize missing combinations.

(b) Ensemble methods for dealing with the details for output classes.

⇒ Ensemble methods such as blending, random forest, etc., can be used to classify attributes by combining the predictions of multiple models. This method avoids the need to create a single model with a large number of output classes.

2、

(1) Training:

To classify the 20 classes in the dataset, transfer learning can be applied by using the pre-trained Alexnet which was trained on a large dataset with 1000 classes. Instead of training a new model, the last fully connected layer of AlexNet can be removed and replaced with a new fully connected layer consisting of 20 units.

(2) loss function:

For face verification, the contrastive loss is a good choice as it designed to reduce the distance between the feature embeddings of two images of same person. while increasing the distance between two images of different people.

(3) Training:

During training process, pairs of face images (one positive pair with images of the same person and one negative pair the different people) will be fed into the network. The network will learn to generate embeddings for each image, which are lower-dimensional representations of the image features.

(4) Inference:

To determine if two face images belong to the same person, I'll compare the distance between their embeddings. By setting a threshold, I can make the classification decision.

3. ACE method is used for separating mixed signals into independent components. To prevent overfitting during the training of the model, we include regularization terms in the overall loss function. In general, we add a constraint to ensure that the output of the interfering category is small, which prevents the model from relying solely on the input data to produce the output. However, if we simply force all the indp. component outputs to zero, we may not able to get the optimal solution that seperates the mixed signal into indp. components.

This is the main goal of ACE, and without regularization term, the model may perform well on training data but poorly on testing data because of overfitting.

4. The first layer is fully connected layer, and input vector is 100.

$\#$ of param $= 4 \times 4 \times 1024 \times 100 + 4 \times 4 \times 1024 = 16446000$

Next 4 layers are fractionally-stride conv layers

Each layer has kernal size 5×5 and stride of 2.

$$\begin{aligned} \# \text{ of param} &= [5 \times 5 \times 1024 \times 512 + 512] + [5 \times 5 \times 512 \times 256 + 256] \\ &+ [5 \times 5 \times 128 \times 128 + 128] + [5 \times 5 \times 128 \times 3 + 3] \\ &= 17213699 \end{aligned}$$

total $= 16446000 + 17213699 = 33629699$ #.

5、

(1) Data collection:

To train a de-fogging system, it's necessary to gather images captured in varias weather conditions. We can take photos in different foggy weather conditions, and clear images can be captured of the same scene on a clear day.

(2) Data Pre-processing:

After collecting data, pre-processing should be done before training. This may includes resizing, cropping and normalizing to ensure that every image having same size and intensity.

(3) Training Procedure:

Using a DL model to learn the mapping between foggy and clear images. There exists lots of different architectures and different loss fuction, which can be used to train the model.

(4) Inference procedure:

Once the model is done, we can apply it to remove fog from foggy images, which turn out to be clear images