

Title: Identity authentication algorithm under complex service scenarios (telecommunications, banking, digital content operators, municipal service and other counter services)

Industrial Applications: □Intelligent Manufacturing □Intelligent Driving □Intelligent Life □Smart Medicine ■Smart City

[Overall background]

With the information era coming, whether for the bank, the telecom operator, or for the government, the operation for the customers is more and more dependent on the information resources and the information network. While the network brings benefits, value and convenience, it also brings huge risk and hidden danger. Many criminals try to take use of false identity information or be anonymous to carry out related business management, to build a tool for illegal crimes, which damage social stability

The crime of telecommunications, finance and network fraud will lead to a social public nuisance, which seriously affects the legitimate rights and interests of the people and destroys the social harmony and stability. It should be more effective to prevent and accurately combat the crime of telecom network fraud, to ensure the legitimate rights and interests of the people, to further improve the business environment and to implement the registration system of the users' real identity information. It can also greatly reduce all kinds of information harassment and illegal and criminal activities caused by anonymity.

[Business background]

Many departments in operators, banks and municipal services, carry out various types of business (including different security levels) for different and large number of user groups every time. The situation is complicated and changeable, which is easy to be drilled by criminals. In order to ensure the security and efficiency of business, it is necessary to verify the authenticity of the user's identity, which requires a variety of verification and authentication techniques for the related video, pictures, sound and other information, and to ensure safety as far as possible.

At the same time, it also needs to meet the requirements of business specifications, high accuracy, high efficiency, and so on. Therefore, many technical capabilities are required to meet the above abilities to face the complex business, the safety grade business, the business acceptance in the unexpected situation, and the related situations in the relevant time and place. Business acceptance, as far as possible to reduce the amount of manual participation, in order to improve the security, but also to save a large number of human and material resources, with good social and economic benefits, so it is imperative to develop a safe identity authentication system with multi biological characteristics.

[Problem description]

In this project, the multi biometric security authentication system under the complex business scene is mainly based on the two most mature biological features of the face identity authentication and sound feature authentication. The face identity authentication mainly through the living body, the picture service standard detection, the counterfeit photo recognition, the similar photo transact, the witness comparison, face search (black and white list detection, etc.)--totally 6 technical methods, are implemented. Sound identity authentication is realized by 2

methods of voice content recognition and sound pattern recognition.

[User expectations]

(1) At present, after the service is accepted, users should be equipped with more staff to check the identity information (including the signature of the business, the authenticity of the picture, the identity information of the staff in the picture, etc.). In face identity authentication, for fake photos, photos not conforming to business standard black and white list detection and so on, it needs a strong ability of "checking all". That is, the system can find the relevant data as far as possible in the face of erroneous data.

(2) "Checking accuracy" of face recognition has the advantage of high reliability. The system can accurately find the data needed in the face of certain sample data.

(3) Application scenarios of the system can be diversified, and can be applied to different business scenarios, to save the workload of the business department and improve the work efficiency.

[Expected economic effect]

(1) Direct economic effect.

It is becoming more and more important for the government, telecom operators, banks and other units to ensure the users' identity security. Once the system is popularized in a large area, the market prospect in China is expected to reach more than 5 billion Yuan (RMB).

(2) Indirect economic effect.

After the system is online, it can effectively avoid the crime of criminals and greatly reduce the incidence of fraud crimes such as telecommunications and finance, and can save billions of losses for the country and the common people every year, contributing to the harmonious and stable society.

[Set up 3 sub items and 1 comprehensive item.]

3 sub items and 1 comprehensive items are set up as follows:

Image class:

The first major category is anti-counterfeiting and fraudulent, which includes 4 sub aspects, namely, living objects, standardized detection of picture business, counterfeit photo recognition, and similar photo processing. The score and weight of each case are shown in Part 5.1 of the competition. The highest score is the first prize.

Second category: identification classes, including witness comparison and face search. The score and weight of each case are shown in Part 5.2 of the competition. The highest score is the first prize.

Speech class:

Voice authentication: speech content recognition and voiceprint recognition. The score and weight of each case are shown in Part 5.3 of the competition. The highest score is the first prize.

Comprehensive class:

After all 3 sub items in Image and classes are summarized. The score and weight of each case are shown in Part 5.4 of the competition. The highest score is the first prize.

[Test description]

1. Organizer will provide relevant test data for each participant team according to each technical method.

2. Organizer will provide 50-500 different test data (similar to final match test data) for 4 cases including living body, picture service standard detection, counterfeit photo identification, and similar photo processing. etc. (similar to the final match test data, the same below) for testing. Data provision will desensitize sensitive data and put it online for reference. For human comparison, face search will provide test data ranging from 1000-10000. The sound data provides 50-100 minutes of single-channel recording data segments and corresponding annotation texts for 50-100 minutes, including quiet and with some noise, except for the normal situation of normal speech speed and regular cavity. In addition, it also includes situations where the tongue and tongue are not standard, slightly dialects, and fast speech speeds.

3. As relevant test data relates to enterprise production information, related parties need to sign a confidentiality agreement. During the test process, the party cannot copy, take test data, etc. to disclose the information provided by the data provision. Otherwise, the participant will be deleted and the organizer will handle it with relevant regulations.

[Game description]

1, Due to the data and information involved in test makers, they will provide the machine environment, including cpu, gpu environment windows, linux2 operating system machines

2. Provide data of similar test data ranging from 50-800 for live games, inconformity with business specifications, image detection, forgery photo identification, and similar photo processing. For match comparison, face search, it will provide 10000-300000 data for competition.

3. For each processing speed of the game image data, such as in the cpu environment, it cannot exceed 3 seconds per picture; in gpu environment it cannot exceed 2 seconds per picture.

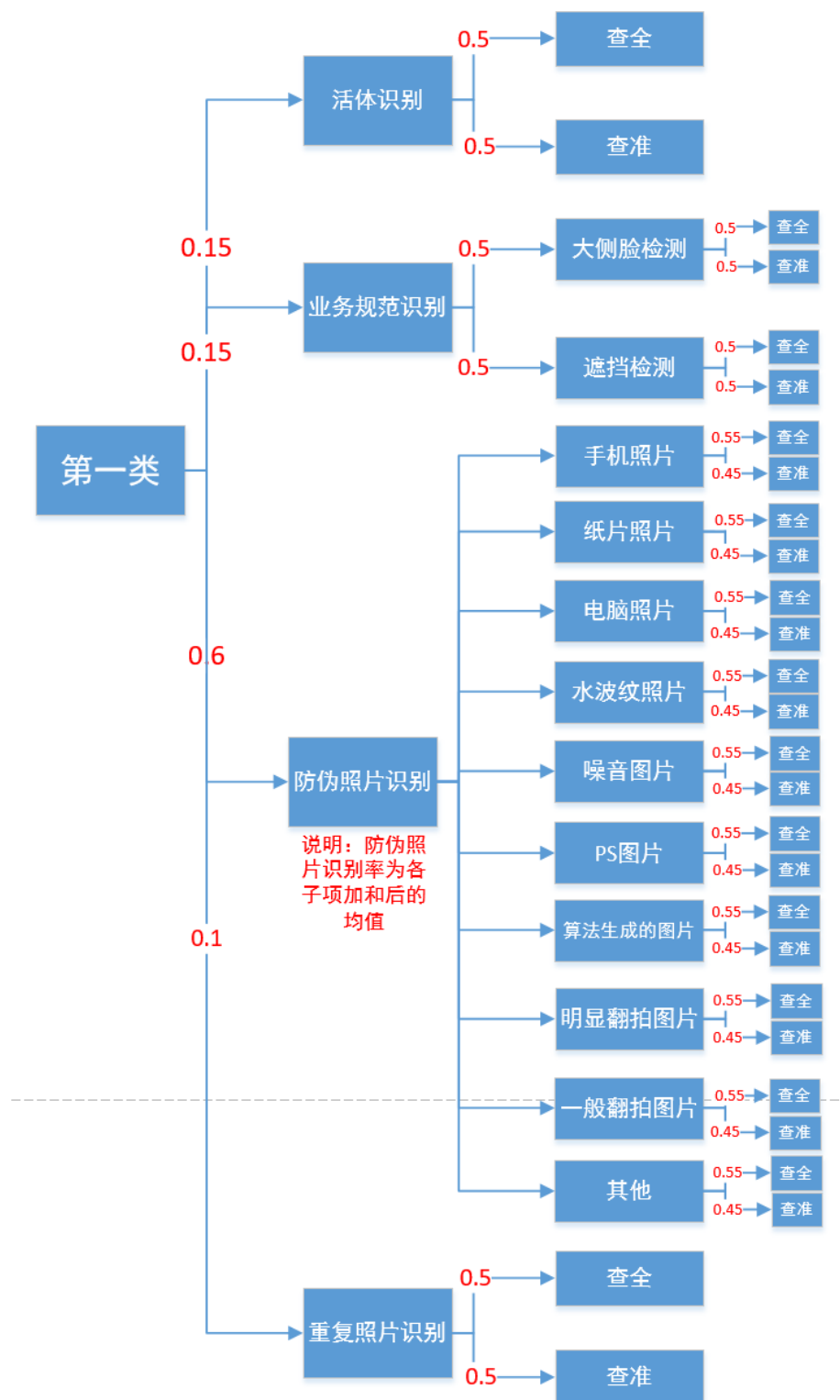
4. For the voice recognition, it should provide 100-200 people 30 seconds -1 minutes of single-channel recordings as competition data (for details, see the relevant test data described above).

5. Score composition on several items (Need re-determining)

5.1. The security-related anti-counterfeit categories include four conditions including living body, picture business standardization detection, forgery photo identification, and similar photo processing. If the items contain sub-items, except for the forged photographs in the sub-items (the sub-items are checked and checked according to the weights of 0.55 and 0.45, the scores of each sub-item) are related to the check and check according to 0.5 and 0.5. The weighted average is used as the percentage of accuracy of the last sub-item, and the average of the sub-items is the percentage of the sub-item.

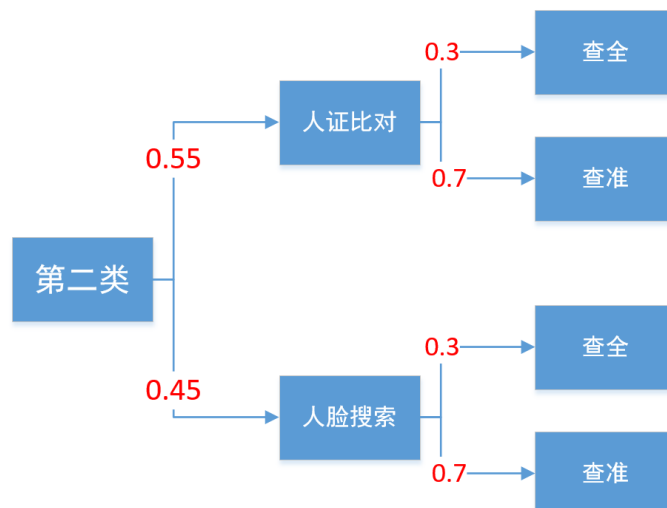
After the score, the scores of the above four items are weighted by the weighted average of 0.15, 0.15, 0.6, 0.1, etc., respectively, and the total score of the large item is 100 points.

As shown below:



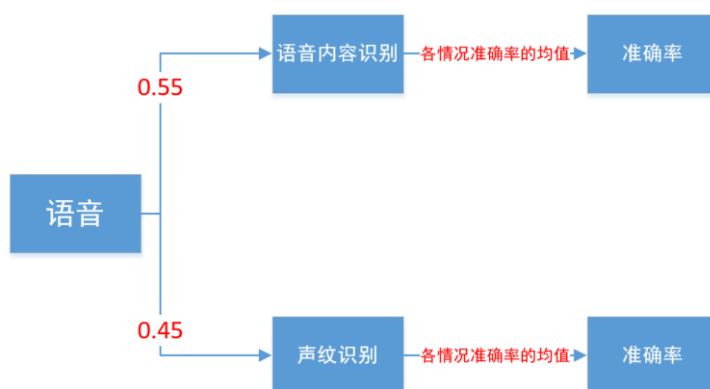
5.2. Involving person identification comparisons, face searches, etc. will be calculated as a percentage ratio according to the first accuracy rate, converted into the corresponding number of scores, and the final score of the large item, total score calculated based on the weighted average of 0.55, 0.45. The total score is 100 points.

As shown below:



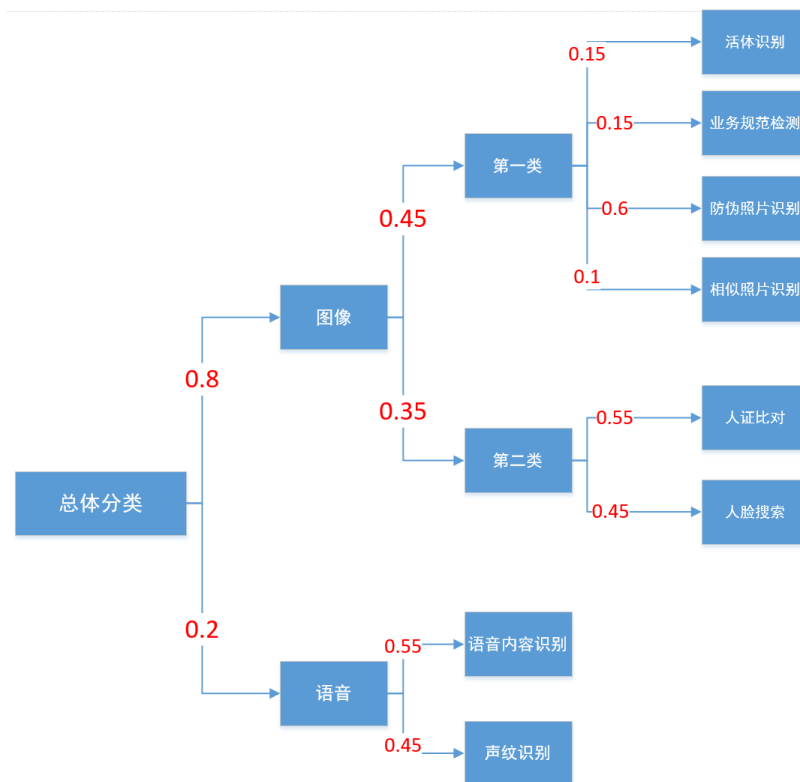
5.3. In terms of voice identity authentication, two categories, i.e., identification content and voiceprint, shall be calculated as a percentage of the first accuracy rate, converted into a score corresponding to the number, and weighted by the weighted average of 0.55 and 0.45 to calculate the last major item. The total score is 100 points.

As shown below:



5.4. The final total score will be weighted by the weights of 0.45, 0.35, and 0.2 for the respective large items (in the above order). The total score is 100 points.

As shown below:



5.5 Final ranking of all participating units according to the percentage system.

[Individual research and development instructions]

To ensure fairness and fairness, mature third-party interfaces and sdk must not be used to promote research, development, and application of artificial intelligence algorithms. The organizer commissioned a team of experts to ensure that the intellectual Whether to conduct self-inspection, the following is the details:

1. Each function of the participating units is required to be independently developed. No third-party interface, sdk (such as face and voice recognition, etc. available on the market) may be invoked. If so, the organizer will reduce or cancel the correspondence according to the situation.
2. In order to protect the intellectual property rights of each participating unit and at the same time, relevant participating units are required to prove their own algorithmic autonomy. If necessary, they can make their own docker images in advance and can be deleted after the competition.
3. Participants are required to explain their own algorithm principles and processing procedures. If there is privacy involved in the participating units, they can be explained in the form of verbal instructions.
4. In order to verify the autonomy of the algorithm entering the ranking unit, the organizer asks relevant experts to verify the autonomy of the code according to the situation, and the mode is determined according to the entry situation. For example, on a specific machine, the participant is asked to use a docker image to compile the code, and then perform the same results for the same data on the same data. At the same time, in order to protect the intellectual property rights of the contestants, the supplier of the machine can participate according to the participants. The requirement to remove the mirror or even redo the entire system.

[Technical path]

Multi-biological features:

- (1) Face identity authentication: biometric verification, inconformity with business specification image detection, various types of forgery photo detection, similar photo identification, human identification comparison, face search;
- (2) Analysis of voice features: speech content recognition and voiceprint recognition.

[Technical content and evaluation index]

1. Live recognition

Content Description: There are 6 actions required: head up, down, left, right, open eyes, open mouth, 2-4 kinds of actions are randomly provided for each recognition.

Evaluation index:

Checking all: Not to miss the action.

Checking accuracy: Accuracy of motion recognition.

Score for Item: Calculate the final score of this item by adding 0.5 weights.

2. Normative inspection of picture services

Content Description

- a) Large-side face detection: referring to face that may be too large, it cannot accurately and completely identify facial features and other key facial information, including the side of the face contour, side face behind the eyes, mouth, etc. that cannot be fully displayed,
- b) Cover detection: facial features can be partially blocked and face information cannot be completely identified, including but not limited to hands, dark glasses, hooded masks, masks, and the occlusion of other objects (business computers, etc.)

Evaluation index

Checking all: Not to miss the above violations.

(Note: The search should be based on the actual situation of the picture. If missed, it will be counted in the check. Other cases are comprehensive, they should be based on the facts, etc. The data provider provide or not few complex images with multiple conditions (including multiple non-conforming business rules, anti-counterfeiting subclasses, etc.) at the time of the game to facilitate the calculation of relevant indicators. The same applies below:

Checking accuracy: The accuracy of detecting violations

Score for Item: Calculate the final score of this item by adding 0.5 weights.

3. Forged photo recognition

It mainly includes the use of various tools for portrait remakes, p-picture synthesis (including photocopying, photocopying, etc.), and requires the following types of forgery, in the case of normal photographs and various forged photographs. Normal photos can be distinguished

- a) Mobile Photo: Recognize photos taken from the phone as the subject
- b) Photographs of paper: Recognize photos with photos on paper as subjects
- c) Computer Photo: Recognizes photos of direct computer use
- d) Water Ripple Photo: Forgery Photographs Identifying Ripple Features
- e) Noise photo: Identified noise (Some people's hair appears similar to a remake Photographs of discharges, etc.)
- f) PS picture: identify PS photos, including shots, headshots
- g) Obviously reconstructed photos: It is not clear that the above conditions are more obvious

Shoot

- h) The image produced by the algorithm: Generate the relevant business management image

through the algorithm and

Faces, such as those generated by GAN related models

i) General remakes of photos: It is unclear how the above conditions are less obvious

Shoot

j) Others: There are conditions including no portraits, celebrity photos, unusual colors, etc.

Evaluation index

Checking all: You cannot miss the above violations. In addition, some fake photographs may show multiple forgery features such as water ripples in the case of mobile phone remakes and even computer remakes. There will be 3 cases, it will be considered that performed by a single feature from the images, as well as multiple features that are also included in this case. And so on in many other cases.

Checking accuracy: Check the accuracy of the violation. In addition, some fake photographs may show multiple forgery features such as water ripples in the case of mobile phone remakes and even computer remakes. There will be three cases. Take this as an example, 0.333 points if one condition that be checked, 0.6666 points if two conditions that be checked, and 1 point if all of the three conditions checked successfully. And so on in many other cases.

In case of analogy

Score for Item: The final score is calculated by adding 0.55, 0.45 weights.

4. similar photo identification

In the process of business management, the use of retained photographs in the case of no change, a slight change, etc., in the case of real people do not have to go to the scene for business, you need to detect the same, similar pictures

Evaluation index

Checking all: Not to miss the violation

Checking accuracy: The accuracy of detecting violations

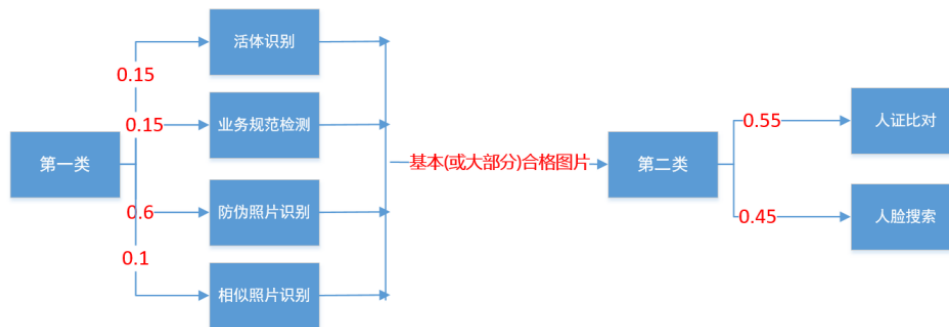
Score for Item: The final score is calculated by adding 0.5 and 0.5 weights.

5. Comparison of witnesses

In the case of complex production of real production, face identity authentication can only be performed after the above-mentioned anti-counterfeiting identification. Therefore, it is required that there should be little or no pseudo-photograph in the theoretical picture of this participation. If there is any, the accuracy of the verification will be counted. , If there are 100 photos involved in certification, search, search, but there are 10 or so forged photos, so check here to enter the next link to 0.9, behind the verification comparison, search the final check The full rate is 0.9* child recall. (Hint: Therefore, if you are interested in participating in the general or participating in this kind of competition, you should improve the entry of qualified pictures.

The rate can be reasonably optimized for various anti-counterfeiting technologies in the past. For example, if you only participate in anti-counterfeiting items, there is no such problem.

The overall process of this major item and the preceding anti-counterfeit items is as follows:



The comparison of human witnesses is performed by comparing the on-site photos with the avatars on the certificates. Then, determine the user's identity

Evaluation index

Check all: You can't miss the handling person (mainly the face can not be recognized). The final recall rate should be multiplied by the security check after the picture enters this item.

Check accuracy: The accuracy is refer to average of accuracy for the same people and different people.

Score for Item: The score is calculated according to the weighted sum of 0.3 and 0.7 respectively.

6. Face search

In real production, only after the above-mentioned anti-counterfeiting identification can the face identity authentication be carried out, so it is required that the pictures involved in this theory theoretically have little or no pseudo-photographs, if any will be counted into the accuracy of the search. The calculation method is multiplied by the search factor at this location. The calculation method is referred to 5

Face search means to capture the user's face information and quickly retrieve the user's information from the database to accurately locate the user's identity, in order to know the black, gray and white list in the business process.

Evaluation index

Checking all: You cannot miss the need to search for people (mainly to identify faces). The calculation method is multiplied by the search factor of the place. The calculation method is calculated by referring to the item search index of 5.

Checking accuracy: Including people as well as strangers on black, gray, white list, the average of the two is the accuracy of the indicator

Score for Item: The final score is calculated by adding 0.3 and 0.7 weights respectively.

7. Voice content recognition

The speech content recognition recognizes the speaker's recording segment, recognizes the speech content, and compares the recognized text with the corresponding annotation speech to obtain the recognition accuracy rate. Voice content recognition should accurately and completely reflect the user's voice information.

Evaluation index

Checking accuracy: Including the accuracy of several cases, such as not, missing, and accurate content. The average value of several is the accuracy rate of the indicator (accurate rate is obtained by comparing the recognized text with the annotation text, including quiet environment and the

accuracy of content with some environments, the average of the two as the accuracy of the indicator)

8. Voice print recognition

The voiceprint recognition achieves the recognition accuracy by comparing the identified user tags with actual user tags. Voiceprint recognition requires accurate identification of the user from the voiceprint information.

Evaluation index

Checking accuracy: Accuracy is that people in the black and white list and in the case of a stranger. The average of the two is the accuracy of the indicator.

[Submission Standard]

Participants are asked to design a set of solutions that meet the above development requirements from the developer's point of view. Participants must independently develop and complete a complete set of systems and cannot introduce third parties to participate.

- (1) Face identity authentication function module (including the above-mentioned several functions);
- (2) Sound identity authentication function module (including the above-mentioned several functions).

[Task list]

- (1) Project documentation, including the principle of the system
- (2) Execution procedure
- (3) A packaged execution environment that contains all (or part of) source code or source code, such as a docker environment with source code.

[reference tool]

None

[References]

None

[Data interface]

None