Assignment: AI-Based Biometric Identification

Type: Internship / Graduation Project

Discipline(s): ICT / Embedded Software / Artificial Intelligence

Level: Bachelor / Master

Assignment description

Can an embedded system identify a person by their face and voice without relying on cloud processing? You will develop an **Edge AI** application on an existing AME hardware platform based on an i.MX processor or similar device, processing image and audio data locally.

The platform includes a dual-camera system with an infrared (IR) camera for 3D depth sensing, a microphone and existing signal processing. Your focus will be on developing the Linux software stack: integrating the image, depth and audio pipelines, selecting suitable AI models and using the processor's neural processing unit (NPU) to accelerate inference.

Together with your supervisor, define how facial and voice-based identification should work individually and in combination, and agree the performance, privacy and security requirements.

Objective

Develop and demonstrate a working Edge AI prototype that identifies enrolled users through facial and voice characteristics and rejects unknown users. Investigate whether combining both biometric methods improves identification reliability, and use depth information to support facial identification and spoofing detection. Run the core AI inference on the NPU and evaluate accuracy, latency and resource use.

Tasks

- **Explore:** Study the hardware, Linux environment, existing camera and audio pipelines, depth capabilities and available NPU tools. Define the use case, operating conditions and acceptance criteria with AME.
- **Select and evaluate AI models:** Compare suitable models for facial detection, facial feature extraction and speaker feature extraction. Assess accuracy, licensing, model size and NPU compatibility. Convert and optimise the selected models, checking the effect of quantisation on accuracy.
- **Build the vision pipeline:** Integrate image acquisition, IR and depth data, preprocessing, AI inference and identity matching. Provide an unknown-user outcome and investigate how depth information can improve robustness and support presentation-attack detection.
- **Build the voice identification pipeline:** Integrate microphone input, speech activity detection, audio preprocessing, speaker feature extraction and identity matching. Evaluate the required speech duration and whether users should speak a fixed phrase or freely. Handle insufficient audio, background noise and unknown speakers.
- **Combine biometric results:** Implement enrollment and removal of facial and voice templates. Compare each method independently with a combined approach. Define how to handle

conflicting identities, uncertain matches, missing inputs and multiple people, including how to associate a voice with the person in view.

- **Use and verify NPU acceleration:** Deploy the selected core AI workloads on the NPU and verify execution through profiling. Document unsupported operations and CPU fallback. Compare CPU and NPU performance, and assess simultaneous vision and audio processing.
- **Protect and test:** Protect biometric templates and enrollment functions, minimise storage of raw images and audio, and agree consent, retention and deletion procedures. Evaluate false matches and false rejections. Test lighting, pose, voice variation, background noise, sensor failures and spoofing attempts such as photographs, recorded speech and synthetic voices. Clearly document detection limitations.
- **Hand over:** Document the architecture, model conversion, installation, configuration and maintenance. Explain the demonstrated capabilities, remaining limitations and recommendations for further development.

Project resources

- Existing AME embedded hardware based on an i.MX or similar processor with an NPU.
- Dual-camera system, including an IR camera for 3D depth sensing.
- Microphone and existing signal processing.
- Access to the platform's Linux development environment and software interfaces.
- Support from embedded software, electronics and AI colleagues, to be agreed with the supervisor.
- Suitable AI tools and consented facial and voice test data, to be arranged during the initial study.

Key deliverables

- Requirements, feasibility assessment and justified model selection.
- Software architecture and a working facial and voice-based identification prototype.
- Integrated image, depth and audio processing pipelines.
- Enrollment, removal and secure management of biometric templates.
- Combined identification logic with defined handling of unknown users and conflicting results.
- NPU-deployed models, conversion scripts and profiling evidence.
- Comparative test report covering individual and combined identification accuracy, performance, robustness and spoofing limitations.
- Source code, build and deployment instructions, maintenance documentation and a final report.