学位報告４

報告番号	※甲　　第　　号

主　論　文　の　要　旨

論文題目　Towards Robust Unconscious Face Recognition for Video Surveillance

（監視映像中で自然に振る舞う人物に対する頑健な顔認識）

氏　　名　張　夢（ZHANG Meng）

論　文　内　容　の　要　旨

The main objective of this thesis is to improve the performance of unconstrained surveillance face recognition quantitatively and qualitatively. Despite the success of deep learning models under constrained face recognition scenarios, the deep features still demonstrate imperfect invariance to wearing a mask, where the whole face image can’t be provided for description. However, the surveillance video provides us with abundant complementary information across frames compared with a single image. Therefore, this research focus on face recognition with wearing face masks and feature aggregation-based face recognition between the multi-fames, and proposes two approaches: firstly, a mask face recognition approach with mask transfer and self-attention, and secondly, a context-aware contribution estimation feature aggregation for surveillance face recognition. Meanwhile, we setup surveillance camera and use our trained models using our proposed research works for access control of laboratory gate to analyze and verify the feasibility of our proposed methods in practical application scenarios.

The first research work presented in this thesis propose a method used for mitigating the negative effects of mask defects on face recognition. Firstly, a low-cost, accurate method of masked face synthesis, i.e., mask transfer, is proposed for data augmentation. Secondly, an attention-aware masked face recognition (AMaskNet) is proposed to improve the performance of masked face recognition, which includes two modules: a feature extractor and a contribution estimator. Therein, the contribution estimator is employed to learn the contribution of the feature elements, thus achieving refined feature representation by simple matrix multiplications. Meanwhile,

the end-to-end training strategy is utilized to optimize the entire model. Finally, a mask-aware similarity matching strategy (MS) is adopted to improve the performance in the inference stage. The experiments show that the proposed method consistently outperforms on three masked face recognition datasets: RMFRD, COX and Public-IvS. Meanwhile, a qualitative analysis experiments using CAM indicate that the contribution learned by AMaskNet is more conducive to masked face recognition.

The second research work presented in this thesis propose a context-aware feature aggregation scheme to aggregate complementary information between different frames. The difficulties in video-based face recognition, such as dramatic pose variations and low quality, can be alleviated by leveraging the rich complementary information between the frames. However, limited by the mini-batch training strategy, the current deep learning methods only utilizes the frames in each batch during training, which ignore the context of the entire video. Therefore, firstly, a two-branch structure is designed as the Context-aware feature Aggregation Network (CAN). Secondly, a context-aware training strategy using a context bank is proposed, which alleviates the limitation of minibatch samples by using the context of the entire video or several images belonging to the same ID and thus achieves global contribution estimation result. Comparative studies on benchmark datasets, such as IJB-C, YouTube Face (YTF), PaSC and COX, confirm that the proposed approach can achieve state-of-the-art level. Meanwhile, qualitative analysis on Multi-PIE dataset indicates that the contribution learned by the CAN is reasonable and beneficial to video face recognition. Based on the above research, a prototype of unconscious face recognition in surveillance scenes under the COVID-19 Pandemic is designed in this thesis to analyze and verify the feasibility of our proposed methods in practical application scenarios.

In summary, this thesis presents methods towards robust face recognition for video surveillance under the COVID-19 Pandemic. Chapter 1 provides introduction, background, research topics and main contributions to this research. Chapter 2 introduce the related work to our research. The proposed methods for mask face recognition and context-aware feature aggregation-based video face recognition are detailed descripted in Chapter 3 and Chapter 4. Finally, chapter 5 concludes this thesis by summarizing our research contributions and provide possible research directions in the future.