

Table 12: Ablation study of data augmentation, cosine learning rate decay, the stride of the second down sampling feature map, regularization parameters L2\_decay and learning rate warm-up for text recognition. Backbone is MobileNetV3\_small\_x0.5. The. number of channel in the head is 48..

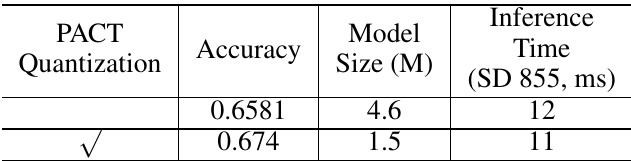


Table 13: Ablation study of PACT quantization for text recognition.

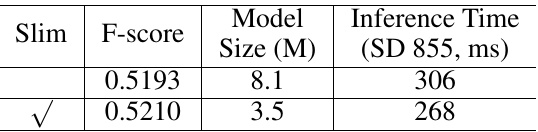


Table 14: Ablation study of the prunner or quantization for the OCR system.

model. We utilize 17.9M training images to learn a text rec- ognizer. Then, use this model as the pre-trained model to fine-tuning the samples for the ablation experiments. When using above pre-trained model, the accuracy will go from 65.81% to 69% and the effect is very obvious.

# 3.5System Performance

Table14shows the ablation study of the prunner or quantiza-. tion for the OCR system. When we use the slim approaches, the model size is reduced 55.7% and the inference time has. accelerated 12.42%. F-score has no impact. The inference time includes pre-process and post-process of each parts of the system. Therefore, FPGM pruner and PACT quantization also are effective strategies for reducing the model size.

also are effective strategies for reducing the model size. To compare the gap between the proposed ultra lightweight OCR system and large-scale OCR system, we also train a large-scale OCR system and use Res18\_vd as the text detector backbone and Res34\_vd as the text recognizer backbone. Table [15] shows the comparison. F-score of the large-scale OCR system is higher than the ultra lightweight. OCR system, but the model size and the inference time of the ultra lightweight system are better obviously..

the ultra lightweight system are better obviously.. Figure13and Figure[14 show some image results of the proposed PP-OCR system for Chinese and English recog-

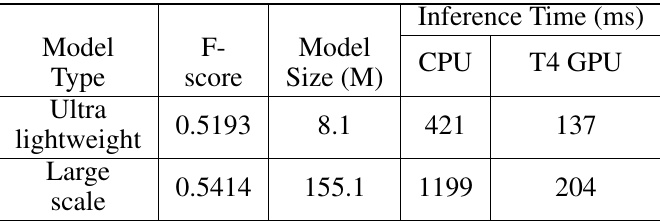


Table 15: Compare between the ultra lightweight OCR sys- tem and the large scale one..

nition. Figure[15 show some image results of the proposed PP-OCR system for multilingual recognition.

# 4Conclusions

In this paper, we propose a practical ultra lightweight OCR system, PP-OCR, which the overall model size is only 3.5M for recognizing 6622 Chinese characters and 2.8M for rec- ognizing 63 alphanumeric symbols. We introduce a bag of strategies to either enhance the model ability or light the model. The corresponding ablation experiments are also provided. Meanwhile, some practical ultra lightweight OCR models are released with a large-scale dataset..

# References

# Chen, P. 2020. GridMask data augmentation. arXiv preprint

Choi, J.; Wang, Z.;, Venkataramani, S.; Chuang, P. I.-J.; Srinivasan, V.; and Gopalakrishnan, K. 2018. Pact: Param- eterized clipping activation for quantized neural networks. arXiv preprint arXiv:1805.06085 .2.2

Cubuk, E. D.; Zoph, B.; Mane, D.;, Vasudevan, V.;, and Le, Q. V. 2019. Autoaugment: Learning augmentation strate- gies from data. In Proceedings of the IEEE conference on computer vision and pattern recognition, 113-123.2.2

Cubuk, E. D.;, Zoph, B.; Shlens, J.;, and Le, Q. V. 2020. Ran- daugment: Practical automated data augmentation with a re- duced search space. In Proceedings of the IEEE/CVF Con- ference on Computer Vision and Pattern Recognition Work- shops, 702-703.2.2