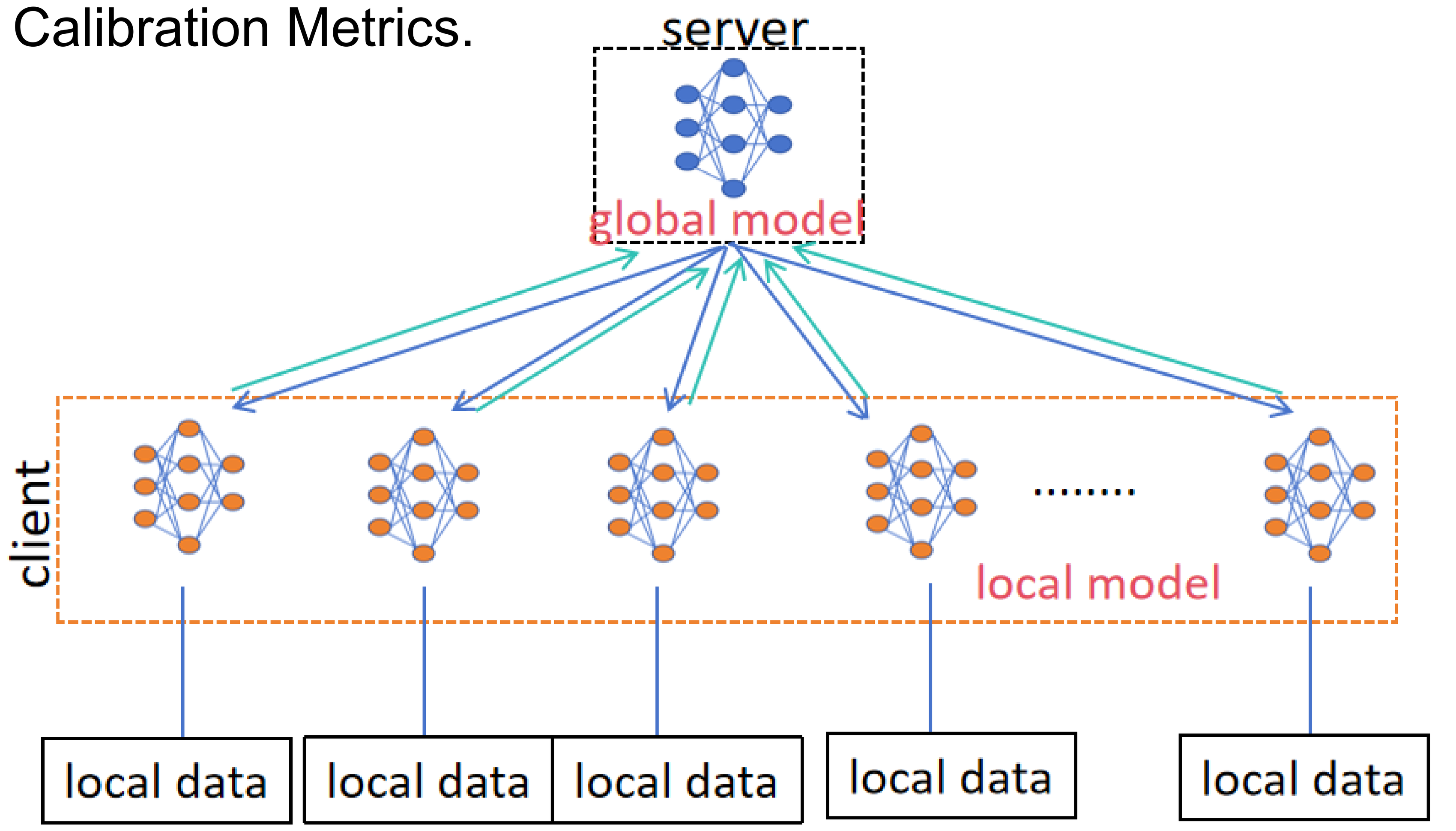
Uncertainty Estimation and Calibration Measurement in Federated Learning for Image Classification

s232251 Bozhi Lyu , s232272 Zekun Wang, s222457 Yichen Jiang

Introduction

- **Federated learning**: a distributed machine learning method aimed at collaborative training on dispersed data sources enhancing data privacy and security.
- **Uncertainty estimation**: quantifying the uncertainty/confidence of predictions. [1]
- **Calibration**: Using metrics to quantify calibration to make predicted probabilities align with the real accuracies. [2]

Keywords: Federated Learning, Uncertainty estimation, Calibration Metrics.



FL Architecture and FedAvg[3]

Algorithm 1 FederatedAveraging (FedAvg). Given by: Global model $w^{(global)}$, Fraction of clients C, Number of clients K, Local epochs E, Local minibatch size B, Maximum rounds $rounds$, Learning rate η, Test Accuracy threshold $acc\_threshold$

Server Do:
$c\_per\_round \leftarrow \max(\text{round}(C \times K), 1)$
for $t = 1$ to $rounds$ **do**
  $clients \leftarrow \text{random.sample}(\{1, \ldots, K\}, c\_per\_round)$
  for each client c in $clients$ in parallel **do**
    $w_{t+1}^{(c)} \leftarrow \text{client\_training}(w_t^{(global)}, E, lr)$
  end for
  $w_{t+1}^{(global)} \leftarrow \frac{1}{C \times K} \sum_{c=1}^{C \times K} w_t^{(c)}$

  if right prediction/total $\geq acc\_threshold$ **then**
    break
  end if
end for

client_training(w, E, η) :
$B \leftarrow (minibatchsize = B)$
for each local iteration e in 1, ..., E **do**
  for each minibatch b in 1, ..., B **do**
    $w_e \leftarrow w_{e-1} - \eta \nabla F(w_{e-1})$
  end for
end for
return w_E

Uncertainty and Calibration Metrics

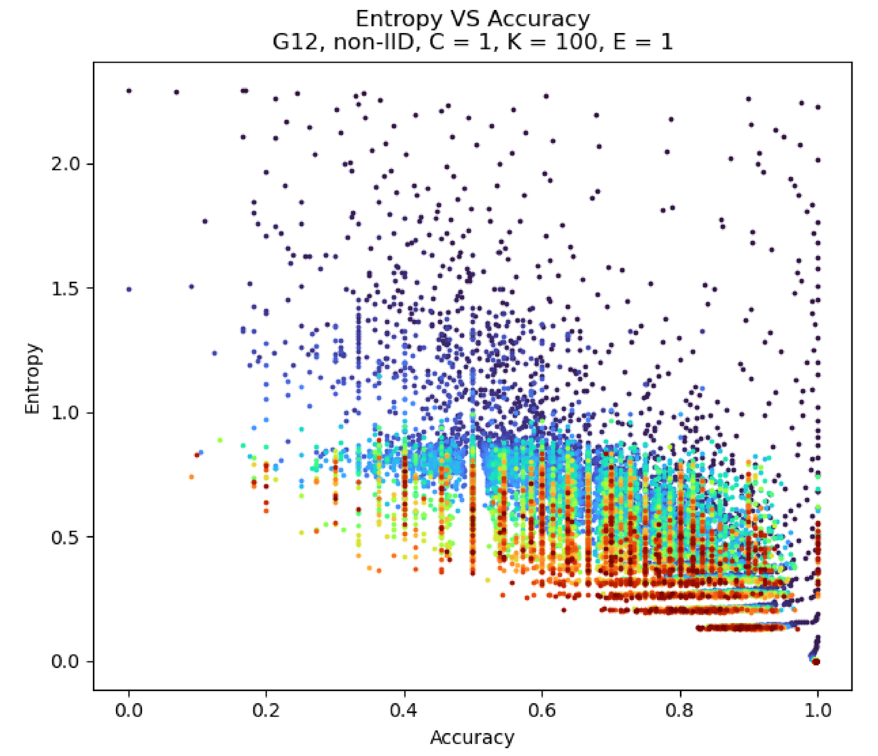
Uncertainty Metrics:

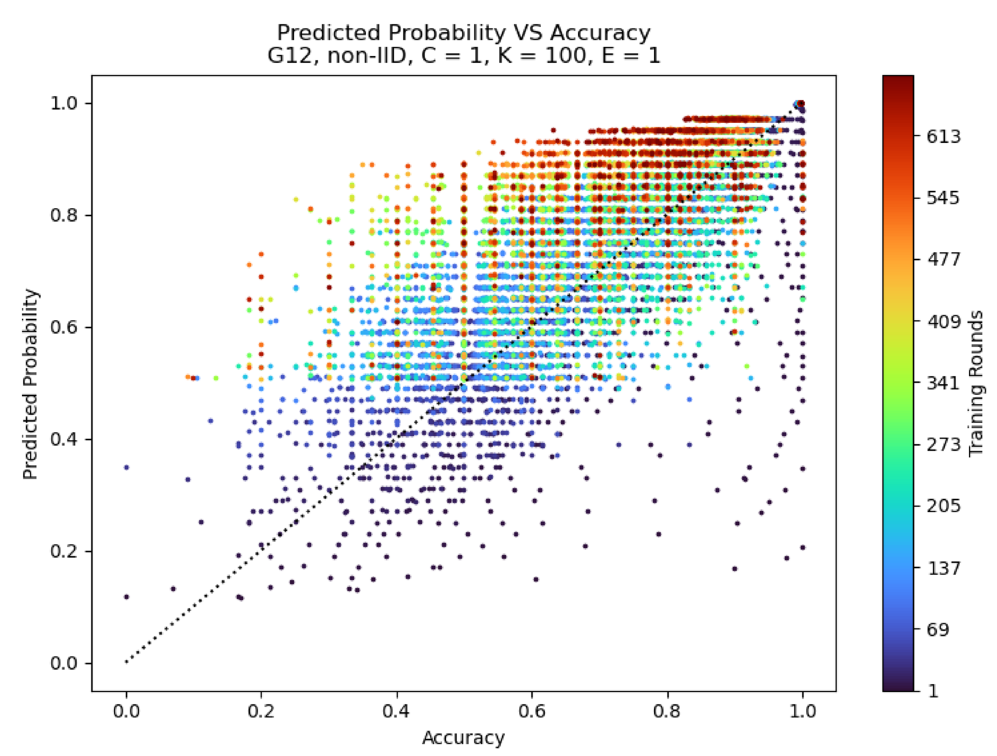
- Entropy: $\mathbb{H}[y|\mathbf{x}, \mathcal{D}_{train}] := -\sum_c p(y = c|\mathbf{x}, \mathcal{D}_{train}) \log p(y = c|\mathbf{x}, \mathcal{D}_{train})$
- Variation Ratios: $\text{variation-ratio}[\mathbf{x}] := 1 - \max_y p(y|x, \mathcal{D}_{\text{train}})$

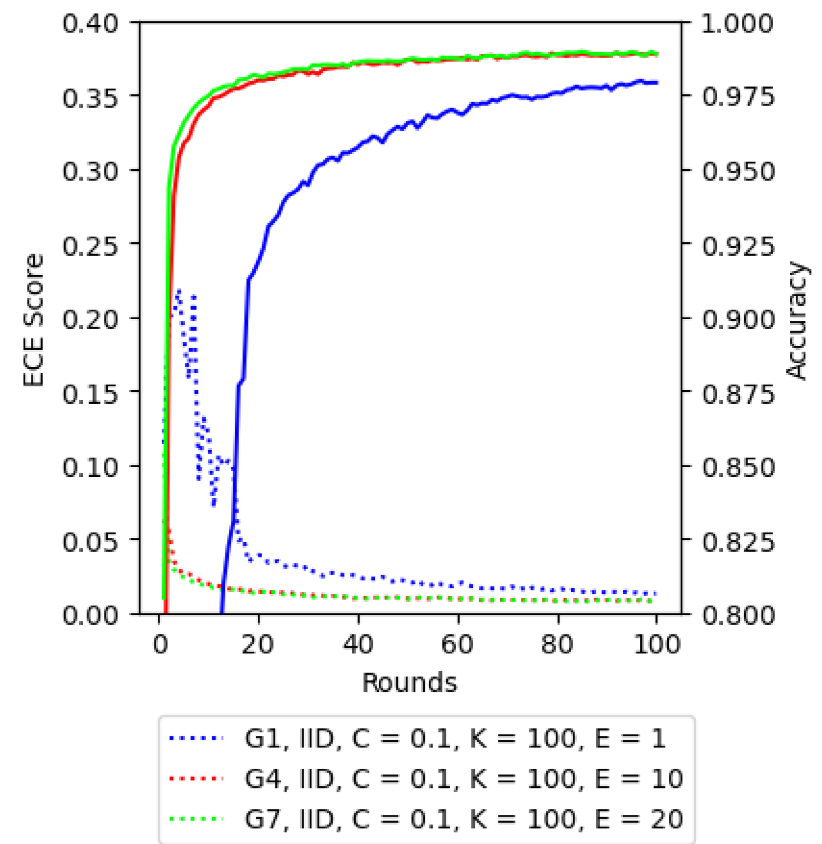
Calibration Metrics: ECE Scores.

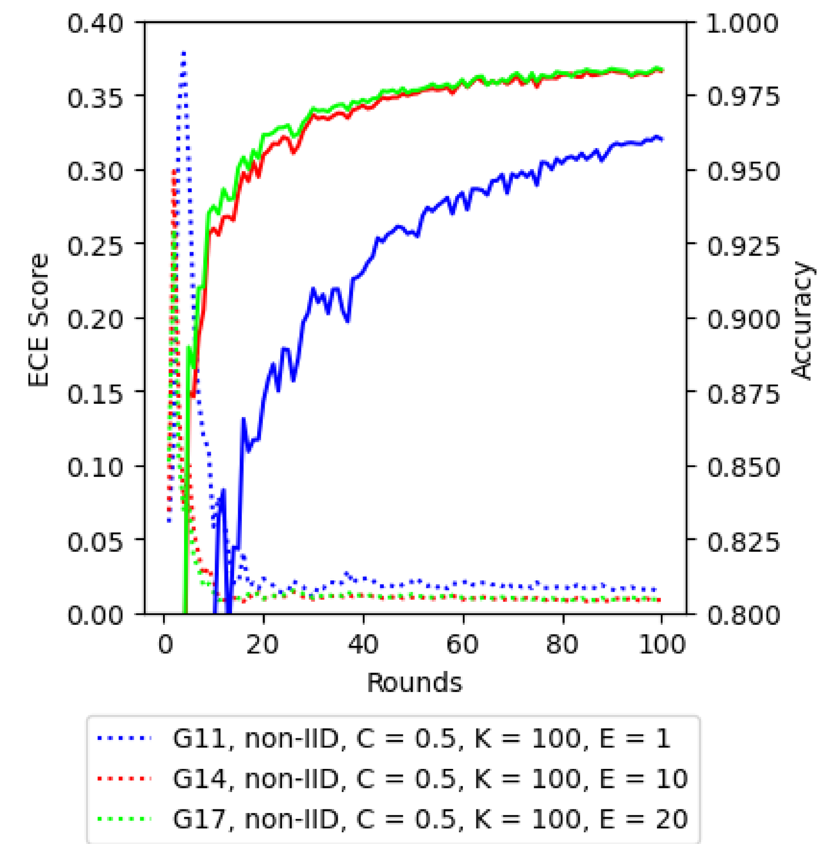
$$\text{acc}(B_m) = \frac{1}{|B_m|} \sum_{i \in B_m} \mathbf{1}(\hat{y}_i = y_i) \qquad \text{ECE} = \sum_{m=1}^{M} \frac{|B_m|}{n} \Big|\text{acc}(B_m) - \text{conf}(B_m)\Big|$$

$$\text{conf}(B_m) = \frac{1}{|B_m|} \sum_{i \in B_m} \hat{p}_i$$

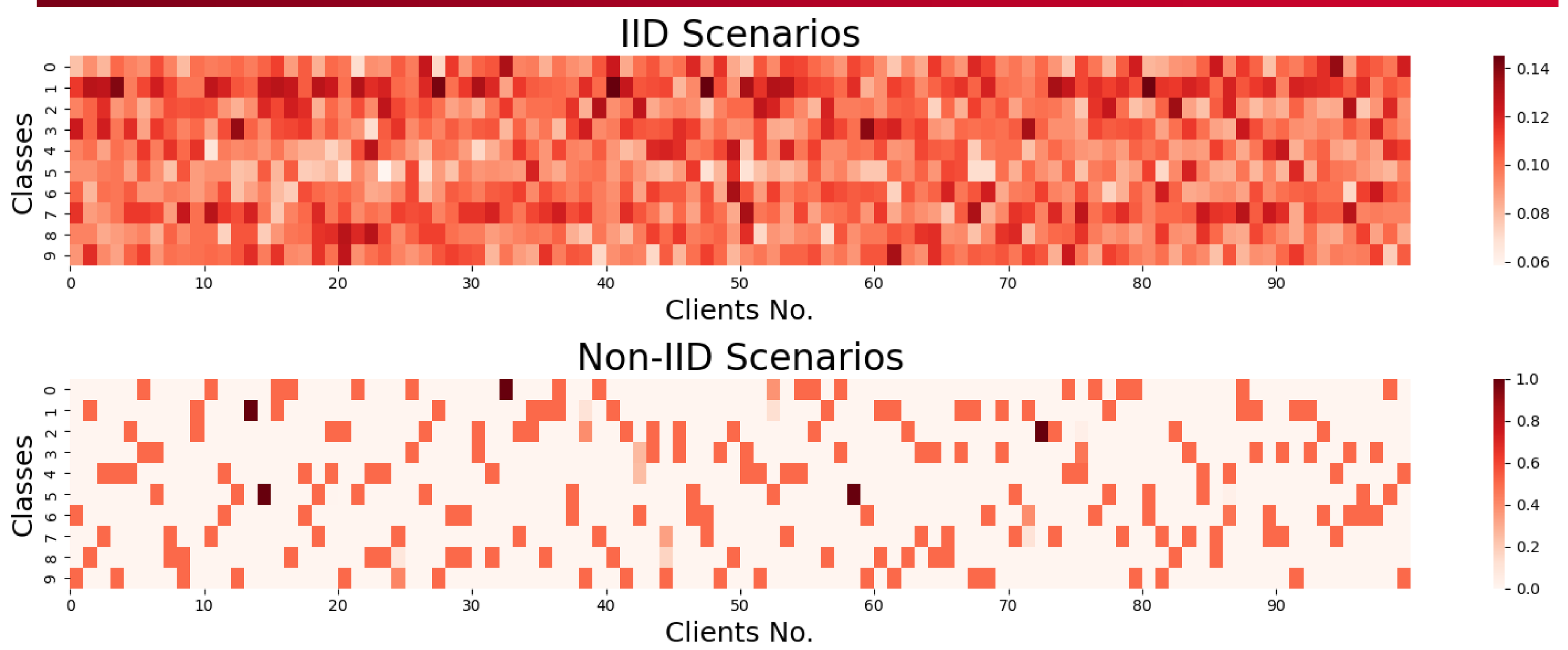








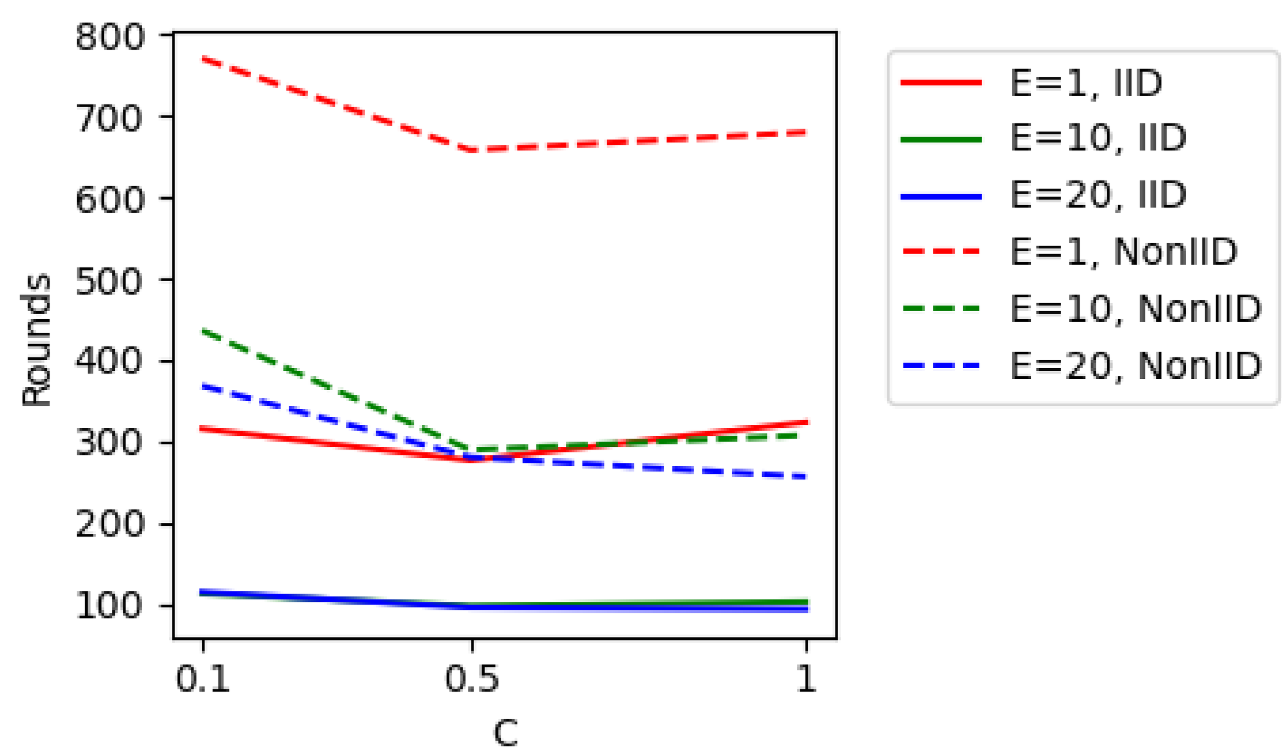
Data Distribution Strategies



Experimental Setup

We used CNN and MNIST dataset. And the target accuracy is set to be 0.99.

- Clients K = 100.
- Local minibatch B = 10.
- Learning rate η = 0.05.
- Different C = 0.1, 0.2, 0.5, 1.
- A set of local epoch E = 1, 10, 20.



Summary

1. **Federated Learning**: We implemented FedAvg algorithm from scratch and reproduced the results of 2nn(mlp) models in Brendan's paper.
2. **Uncertainty Estimation and Calibration measurements**: We implemented entropy and VR as indicators of uncertainty estimation, and ECE score to make calibration quality measurements.
3. We explored the influence of different hyperparameters in FL architecture(CNN) on accuracy, uncertainty and calibration quality.

References

[1] Bo Li [12], Tommy Sonne Alstrøm [2]. On uncertainty estimation in active learning for image segmentation. arXiv:2007.06364v1 [cs.CV] , 2020
[2] Sunil Thulasidasan[1,2], Gopinath Chennupati[1], Jeff Bilmes[2], Tanmoy Bhattacharya[1], Sarah Michalak[1]. On Mixup Training: Improved Calibration and Predictive Uncertainty for Deep Neural Networks. arXiv:1905.11001v5 [stat.ML] , 2020
[3] Jakub Konečný, H. Brendan McMahan, Daniel Ramage, Peter Richtárik. Federated Optimization: Distributed Machine Learning for On-Device Intelligence. arXiv:1610.02527v1 [cs.LG], 2016