

DeepFoodCam : DCNNによる101種類食事認識アプリ

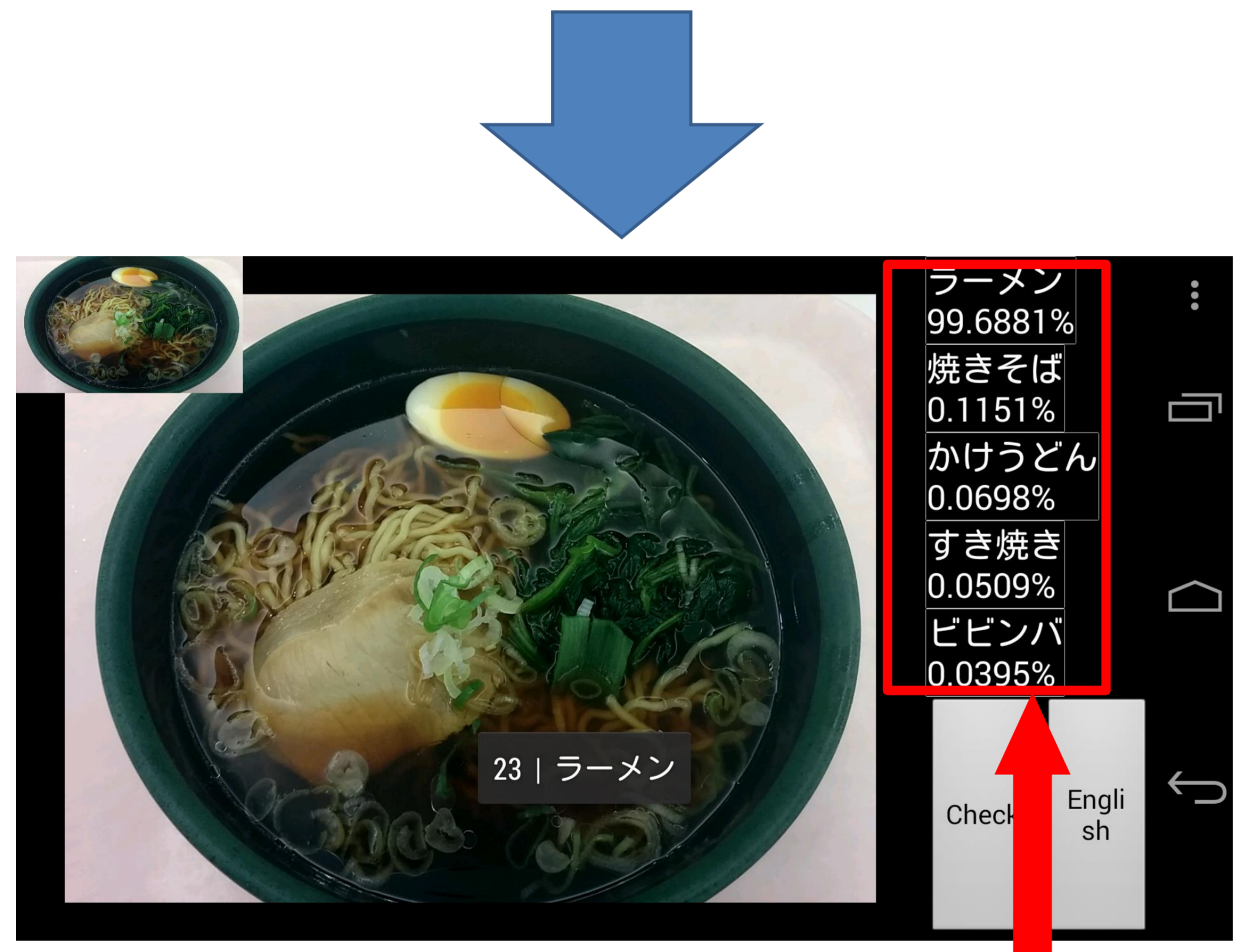
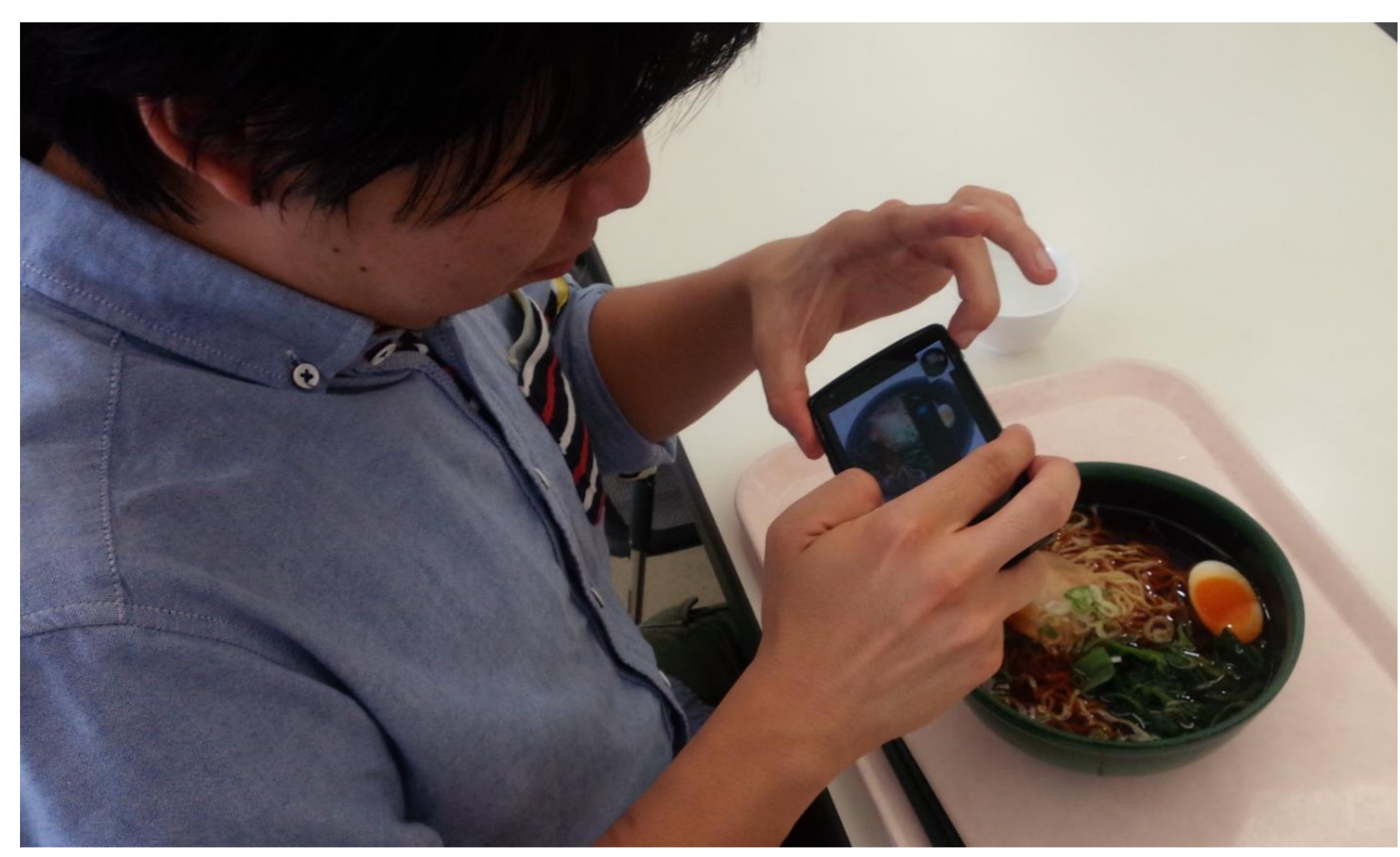
電気通信大学大学院 情報理工学研究科 総合情報学専攻 岡元晃一, 柳井啓司

DeepFoodCam の 特徴

- **スマートフォンアプリで DCNN**
(アプリ内で完結. サーバ不要)
- **101種類**の食事画像認識
- **0.5秒程度**での実行
- 上位5位以内で**93.5%**の精度
- 最大4bitまで **パラメータ圧縮**

DeepFoodCam 概要

使用方法はスマホ撮影するだけ!



認識画像の上位5位までの食事名, 確率

全ての処理はスマートフォン上で実行

- DCNNによる食事認識
- 100種類食事 + 非食事分類
 - 食事/非食事判別率: 98.5%
 - 高速, 低メモリ, アプリ内完結

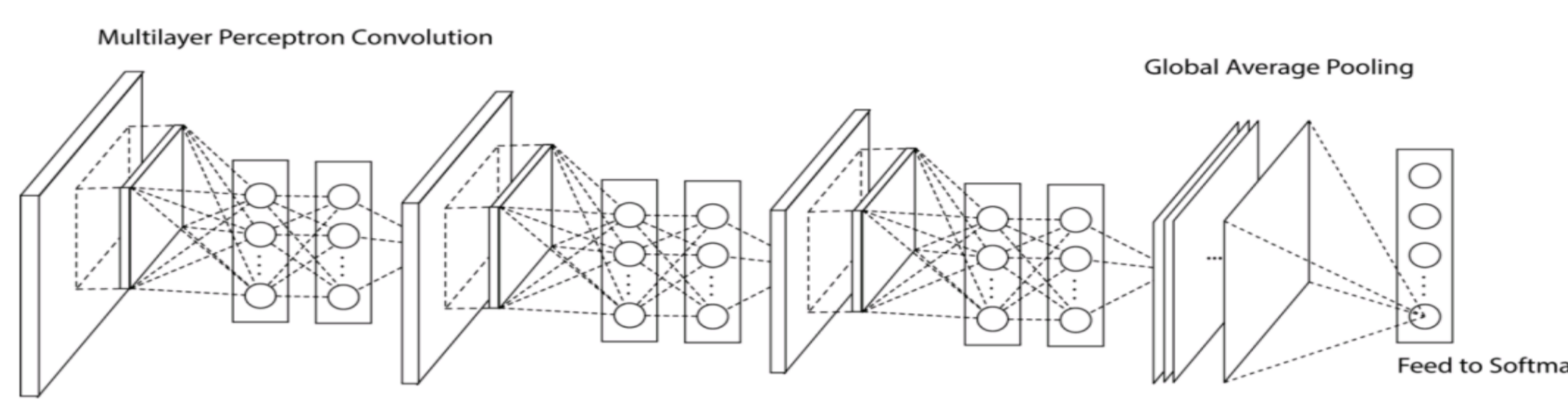
食事認識システム公開中

現状 Android, twitter の2種類
— 現在iOSへの対応も行っている

認識手法: Deep Convolutional Neural Network

モバイルでの実装を考慮しNetwork In Network(NIN)を使用
— パラメータ及び必要メモリの大幅な削減を実現

Network構造	パラメータ数	必要メモリ (32bit)	8bit(圧縮)	4bit
AlexNet	6000万	230M		
NIN	750万	30M	7.5M	3.8M



Network In Network [1]

- コンボリューション層のみ.
- 全結合がない.
- パラメータ数が大幅に少ない.
- 任意画像サイズ対応可能.

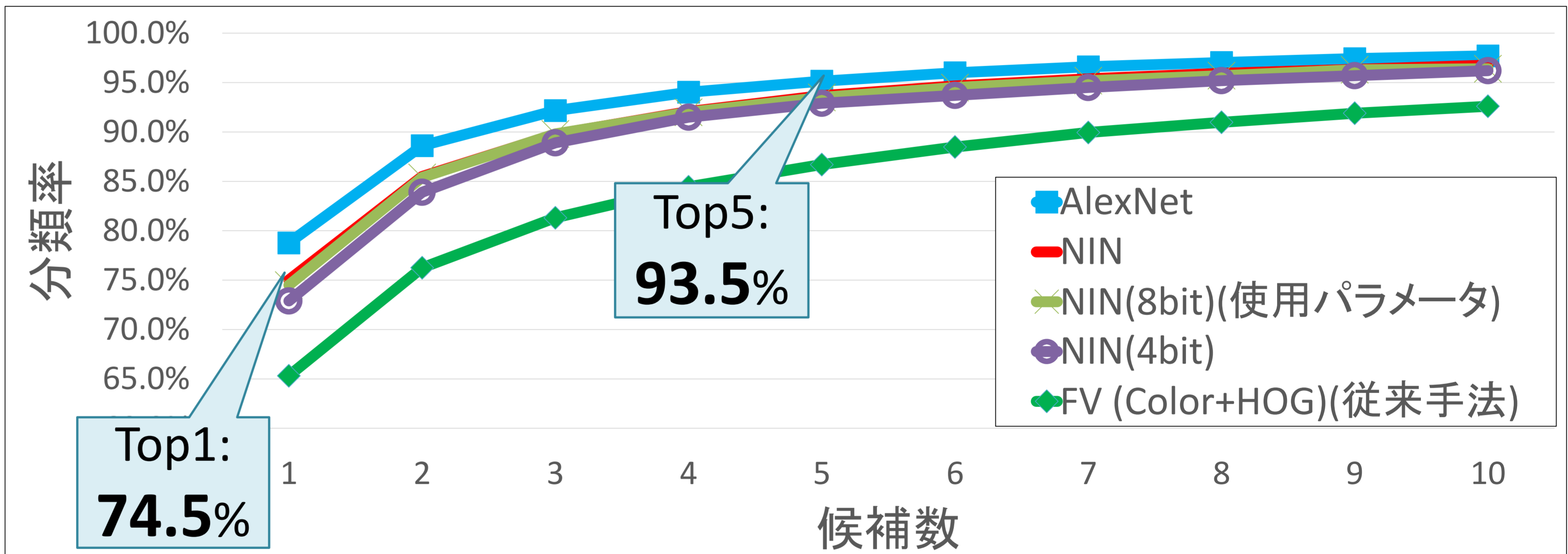
- 2000種類(ILSVRC+Foodの下位語1000)で **プレトレーニング**.
- UEC-Food100 (約1万枚) + 非食事1万枚で **ファインチューニング**.

実装: 高速化のための工夫

- マルチコア対応 (pthread) (クアッドコアの活用) (約3倍)
- Flattened ConvNet [2] の使用 (+30%)
- 活性化信号のスパース性を利用した計算の効率化 (+40%)
- 画像サイズ可変 (224² ⇒ 180²(+40%), 160²(+50%))
- OpenBLAS(SIMD命令)の利用(効果なし) (原因:Neonは2コアのみ)

実験

性能評価 (UEC-Food100 (<http://foodcam.mobi/>) 5-fold CV)



	1	2	3	4	5	6	7	8	9	10
AlexNet	78.8%	88.6%	92.2%	94.0%	95.1%	96.0%	96.6%	97.0%	97.4%	97.7%
NIN	75.0%	85.5%	89.8%	92.1%	93.7%	94.7%	95.4%	95.9%	96.3%	96.8%
NIN(8bit)	74.5%	85.4%	89.8%	92.0%	93.5%	94.5%	95.2%	95.7%	96.3%	96.4%
NIN(4bit)	72.9%	83.9%	88.9%	91.5%	92.9%	93.7%	94.5%	95.2%	95.7%	96.2%
FV(従来手法)	65.3%	76.3%	81.3%	84.5%	86.7%	88.5%	90.0%	91.0%	91.9%	92.6%

実行時間評価

Nexus5 : 2.26GHz(クアッドコア) 1枚の処理 約0.5[sec]

食事認識Androidアプリ

ダウンロード・デモサイト:
<http://foodcam.jp/>

食事認識bot

[@foodimg_bot](https://twitter.com/foodimg_bot)宛に食事画像URLをtweet replyで認識結果が送られてきます

[1] Lin, M., Chen, Q. and Yan, S.: Network In Network, Proc. of International Conference on Learning Representation Conference Track (2014).
[2] Jin, J., Dundar, A. and Culurciello, E.: Flattened Convolutional Neural Networks for Feedforward Acceleration, Proc. of International Conference on Learning Representation Workshop Track (2015).