AR DeepCalorieCam:An iOS App for Food Calorie Estimation with Augmented Reality

Ryosuke Tanno, Takumi Ege and Keiji Yanai
(The University of Electro-Communications, Tokyo)

Background & Objective

Meal management apps enable us to record food calories.
Some of them need human help for calorie estimation.

Image-based food calorie estimation using recipe information with Augmented Reality

[DeepCalorieCam]

[AR DeepCalorieCam]

Method: Multi-task CNN for calorie estimation

Related work ① : Miyazaki et al.[1] 2011

Search-based food calorie estimation with **conventional features**.
Similar image retrieval with SURF and color histograms and so on.
Calculate food calories from retrieved images' calories.

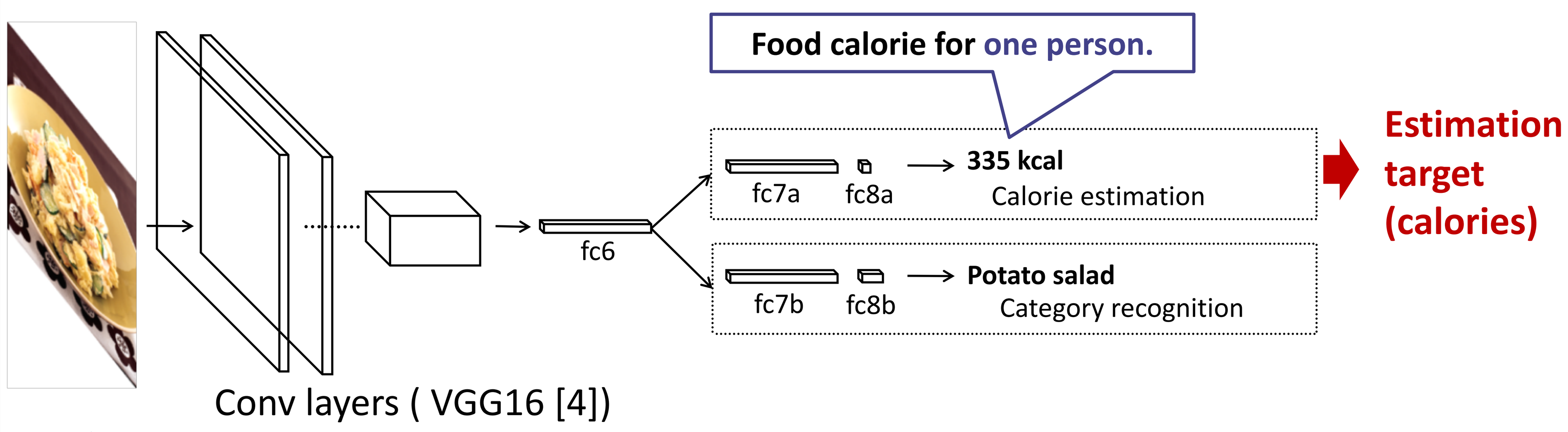
➡ We propose regression-based method using **CNN**.

Related work② : Chen and Ngo[2] 2016

Multi-task estimation of food categories and food ingredients.
Multi-task CNN[3] of food categories and food ingredients
Multi-task learning improve both task's performance.

➡ We use **multi-task CNN for calorie estimation.**

[Our network]



➡ **It is expected to improve the accuracy of each task.**

[Loss function]

(1) Calorie estimation loss: $L_{cal} = \lambda_{re} L_{re} + \lambda_{ab} L_{ab}$ (λ is the weight on the loss.)

(2) Food Category loss: $L_{cat} = -\sum_{k=1}^{n} g_k \log y_k$ (softmax cross entropy)

relative err. loss $L_{re} = \frac{|y-g|}{g}$ absolute err. loss $L_{ab} = |y-g|$

(y is an estimated food calorie. g is ground-truth.)

[DATASET]

For training our network, we collected calorie-annotated food photos from the online cooking recipe sites.

Food **15** categories.
A total of **4877** images.

[Implementation]

(1) Train with **Keras**(backend **TensorFlow**) framework
(2) Convert **Keras** model to **CoreML** model for iOS deployment
(3) Display calorie estimation result using **Apple ARKit** framework

Application Demo:

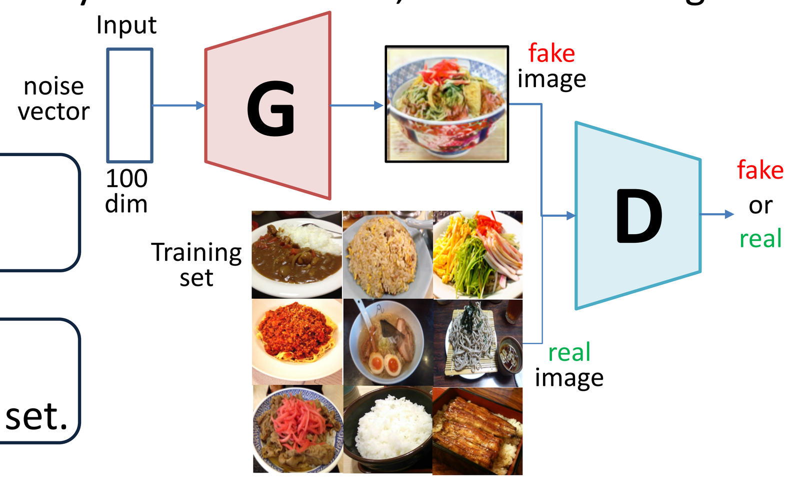
[Calorie Estimation with AR]

[Food Image Transfer using GANs]

My Pet Project Introduction:

Background: Food Image Transfer using Generative Adversarial Networks(GANs)

GANs are a kind generative models designed by Goodfellow et all in 2014.
In a GAN setup, two differentiable functions, represented by neural networks, are locked in a game.
The two players, the **Generator** and the **Discriminator**, have different roles in this framework.



Generator Player 1
The **Generator** tries to produce data that come from some probability distribution.

Discriminator Player 2
The **Discriminator**, acts like a judge. It gets to decide if its input comes from the **G** or from the true training set.

In summary, the game follows with:

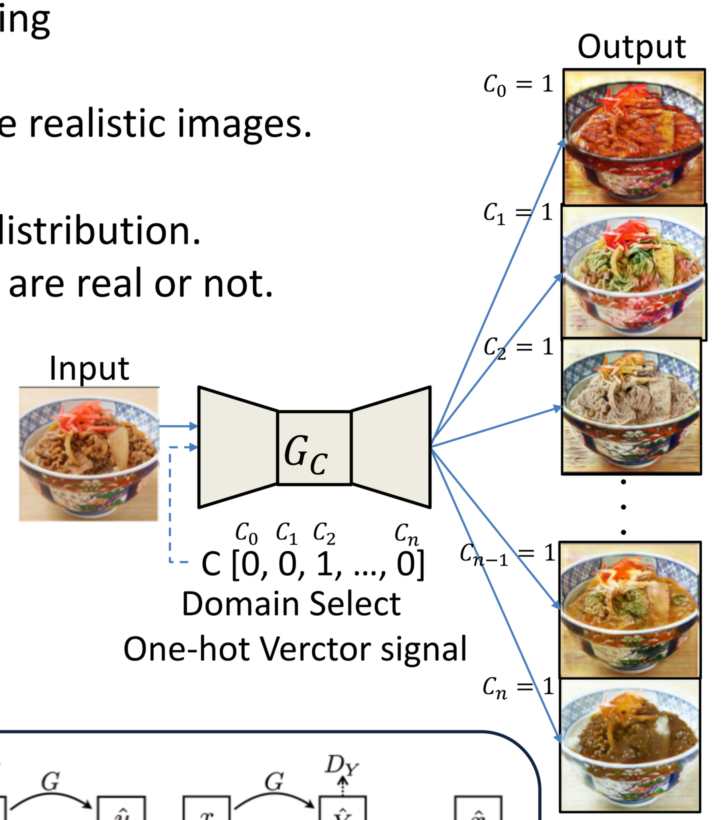
- The **Generator** trying to maximize the probability of making the discriminator mistakes its inputs as real.
- The **Discriminator** guiding the generator to produce more realistic images.

In the perfect equilibrium, as a result,

- The **Generator** would capture the general training data distribution.
- The **Discriminator** is always unsure of whether its inputs are real or not.

Objective:

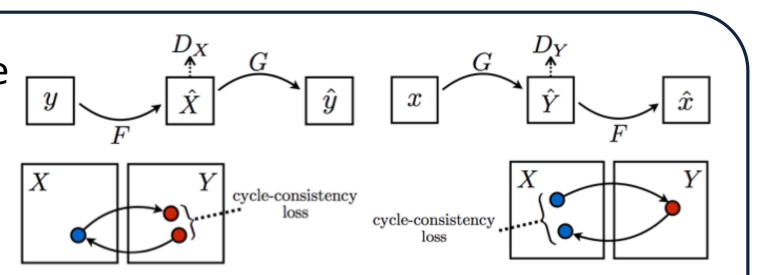
Transfer food images to multiple domains with high quality using the GAN method for dietary images



Method: conditional CycleGAN

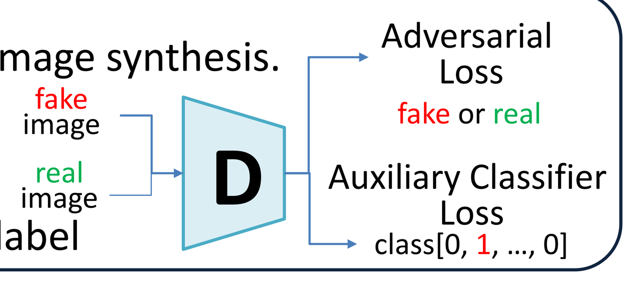
Related work ① : CycleGAN[3] ICCV2017

An approach for learning to translate an image from a source domain X to a target domain Y.
➡ They introduce a **Cycle Consistency Loss** to push F(G(X))≈X (and vice versa).
Cycle Consistency Loss(CCL):
Capture the intuition that if they translate from one domain to the other and back again they should arrive at where they started.

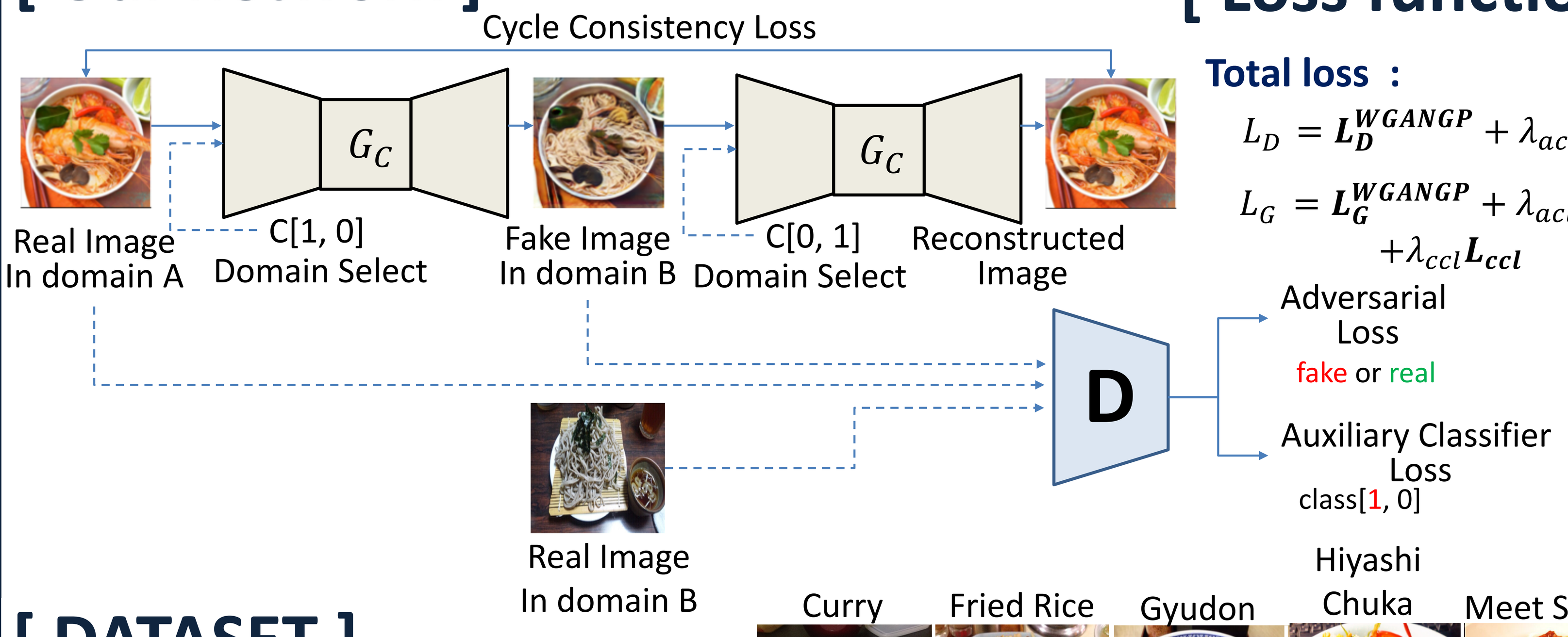


Related work ② : ACGAN[4] ICML2017

New methods for the improved training of GANs for image synthesis.
➡ They introduce a **Auxiliary Classifier Loss** to make high quality image .
Auxiliary Classifier Loss(ACL):
Every generated sample has a corresponding class label



[Our network]



[Loss function]

Total loss :

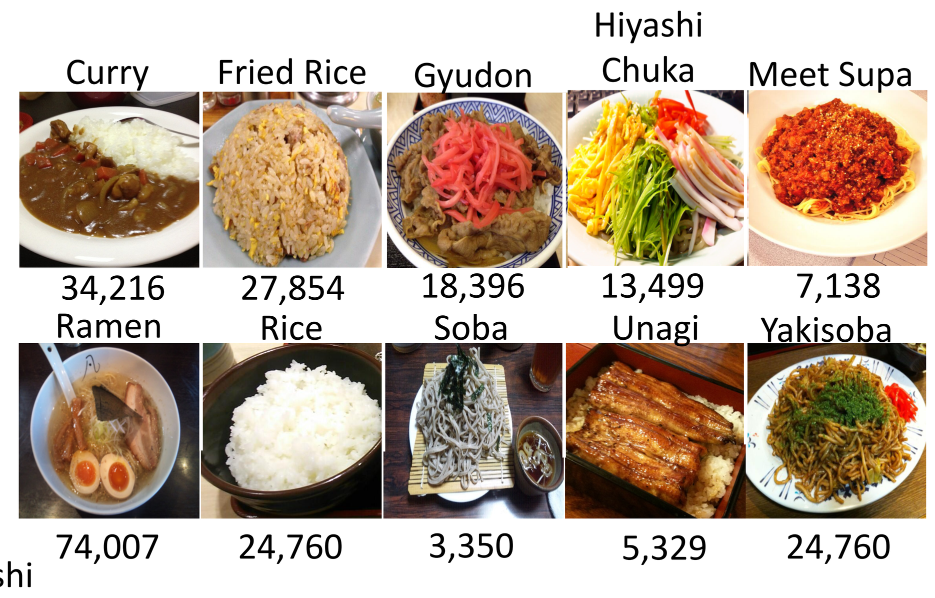
$L_D = L_D^{WGANGP} + \lambda_{acl} L_{acl}^{real}$

$L_G = L_G^{WGANGP} + \lambda_{acl} L_{acl}^{fake} + \lambda_{ccl} L_{ccl}$

[DATASET]

For training our network, we collected 10 category food photos from the Twitter.

Food **10** categories.
A total of **230,053** images.



Result:

Conclusions

[Calorie Estimation with AR]

- We proposed **food calorie estimation app with a multi-task CNN using Augmented Reality.**
- Multi-task learning **improved both food calorie and category estimation.**

[Food Image Transfer using GANs]

- We proposed **food image transfer using conditional CycleGAN.**
- Conditional CycleGan can **convert multiple domains while keeping the shape of the food.**

[1] T. Miyazaki, G. Chaminda, D. Silva, and K. Aizawa. Image - based calorie content estimation for dietary assessment. In Proc. of IEEE ISM Workshop on Multimedia for Cooking and Eating Activities, 2011.
[2] J. J. Chen and C. W. Ngo. Deep-based ingredient recognition for cooking recipe retrival. In Proc. of ACM International Conference Multimedia, 2016.
[3] J. Y. Zhu, T.Park, P. Isola, A.A. Efros, Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In Proc. of IEEE International Conference on Computer Vision, 2017.
[4] A. Odena, C. Olah, and J. Shlens. Conditional Image Synthesis With Auxiliary Classifier GANs. In Proc. of the 34th International Conference on Machine Learning, 2017.
[5] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, Improved Training of Wasserstein GANs. Advances in Neural Information Processing Systems, 2017.