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

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.


We use multi-task CNN for calorie estimation.

[ Our network ]

Food calorie for one person.

Conv layers (VGG16 [4])

It is expected to improve the accuracy of each task.

[ Loss function ]

(1) Calorie estimation loss:

\[ L_{\text{cal}} = \lambda_{w} L_{\text{rec}} + \lambda_{b} L_{\text{ab}} \]

(2) Food Category loss:

\[ L_{\text{cat}} = \sum_{k=1}^{n} g_{k} \log y_{k} \]

(softmax cross entropy)

[ DATASET ]

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

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 of generative models designed by Goodfellow et al 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

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

Discriminator

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.
- 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


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.


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.


[ Loss function ]

(1) Real Image (in domain A)

\[ L_{\text{real}} = \max(0, 1 - \mathbb{E}_{x \sim D_{\text{real}}} [\hat{D}(x)]) + \lambda_{b} \mathbb{E}_{x \sim D_{\text{real}}} [\hat{L}_{\text{ad}}(x)] \]

(2) Generated Image (in domain A)

\[ L_{\text{gen}} = \min(0, 1 - \mathbb{E}_{x \sim D_{\text{gen}}} [\hat{D}(x)]) + \lambda_{b} \mathbb{E}_{x \sim D_{\text{gen}}} [\hat{L}_{\text{ad}}(x)] \]

[ DATASET ]

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

A total of 230,053 images.

Result:

Food category images:

[ 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.