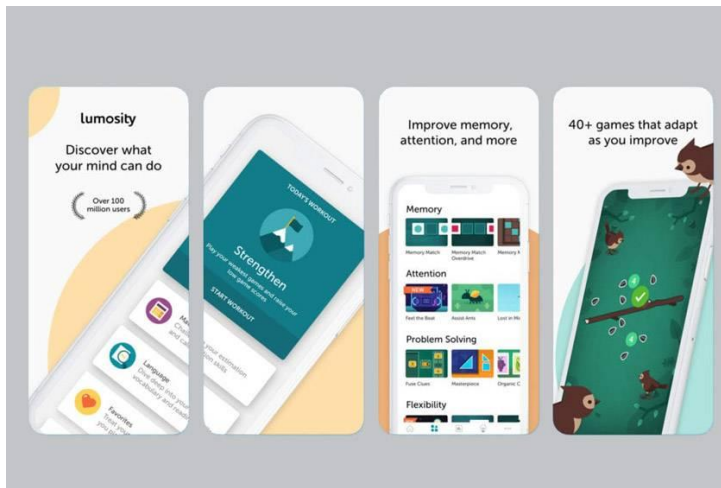
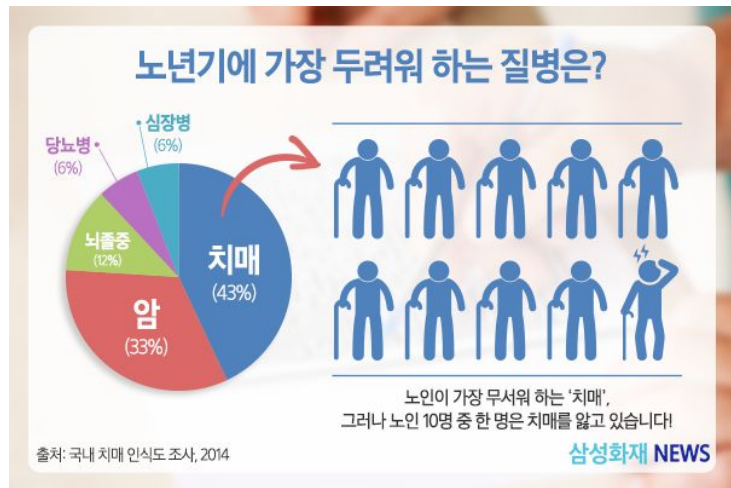


Image Description Application for Dementia Prevention

Seonghae Jo
Mobile Computing and its Application
Mini Project Proposal

Motivation



Goal

- Help the users train their brain
- More interaction with the user and the real world context
(Interact with not only user but also user's context)
- Help the users memorize their daily life

Key Idea



Two children play in the water

1. User captures an image
2. User tries to describe the captured image to train brain
3. ML model generates the caption for the image
4. User compares his/her answer with the caption
5. Do other things (Share with family, Write a diary ...)

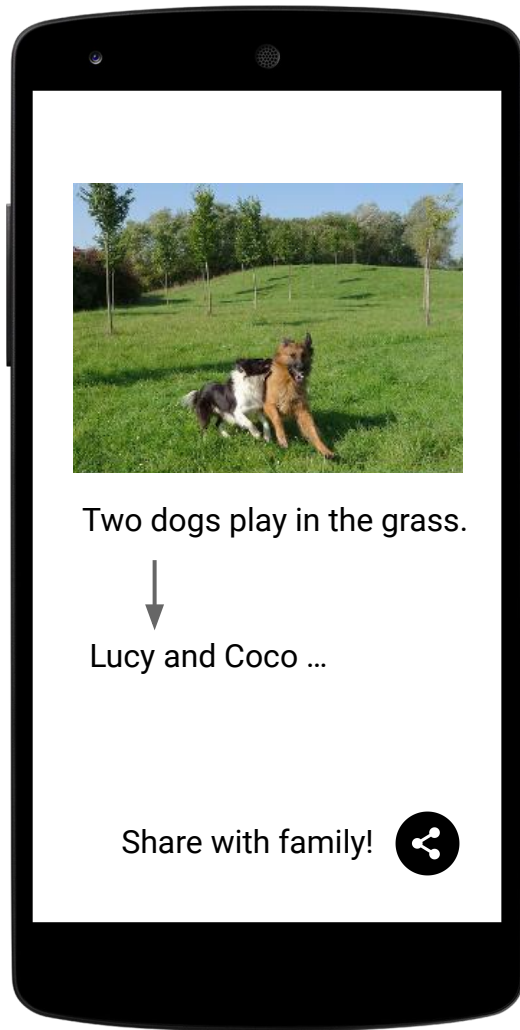
Usage Scenario

Normal mode

Show the caption immediately after capturing the image.

Users might edit the caption.

Users might share with family.



Usage Scenario

Quiz mode

Hide the caption before user interaction.

User trains his/her brain by making sentences.

User can compare his/her answer with ML generated caption.



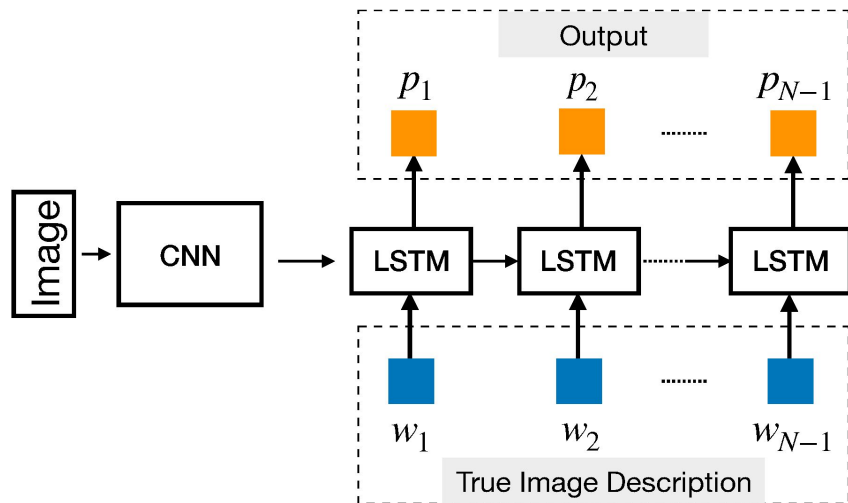
What is this image about?

Your answer:

Two dogs sit on the grass.

Compare with my answer!

Implementation Plan



- Study and select image caption (pre-trained) model
- On edge device, Cloud offloading, or Splitting ?
- Tensorflow Lite Conversion
- Preprocess the captured image (ex. resolution)
- Improve performance

Project Timeline

- (11/19) Study & Choose the image caption model. Determine how to prepare and improve ML model.
- (11/26) Tensorflow Lite Convert & Prepare ML model in Android
- (12/10) Android application development

Related works

- Ning Wang (2022), Efficient Image Captioning for Edge Devices

<https://arxiv.org/abs/2212.08985>

- Oriol Vinyals (2016), Show and Tell: Lessons learned from the 2015 MSCOCO Image Captioning Challenge

<https://arxiv.org/abs/1609.06647>

<https://github.com/KranthiGV/Pretrained-Show-and-Tell-model>