REMOTE VET<>AIML

Akide Liu – Research Assistant, AIML, UofA

Dr. Minh Hieu Phan – Research Fellow, AIML, UofA

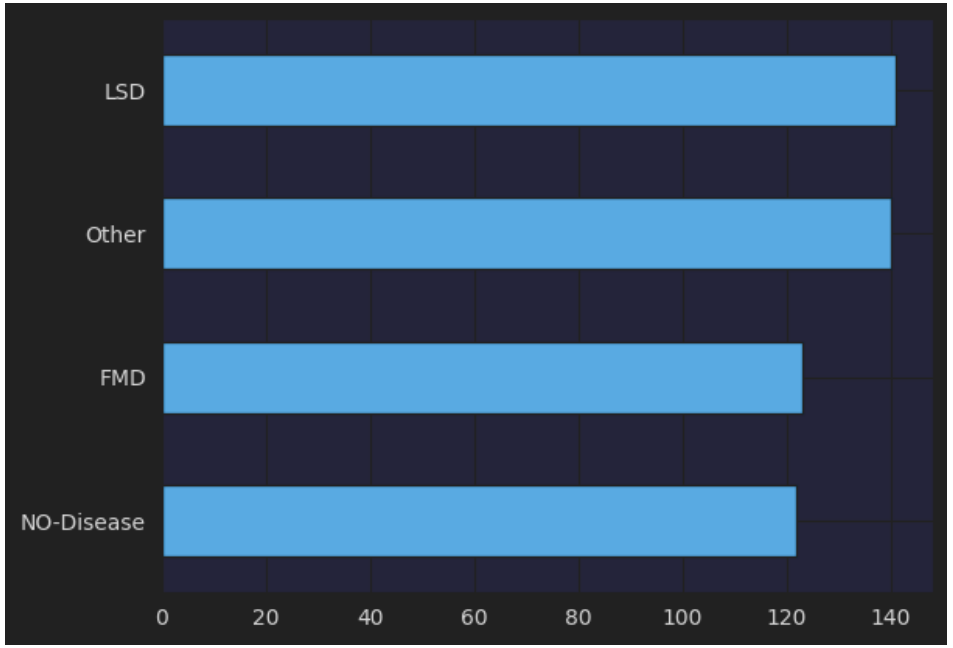
Dr. YiFan Liu – Assistant Professor, AIML, UofA

{akide.liu, vuminhhieu.phan, yifan.liu04}@adelaide.edu.au

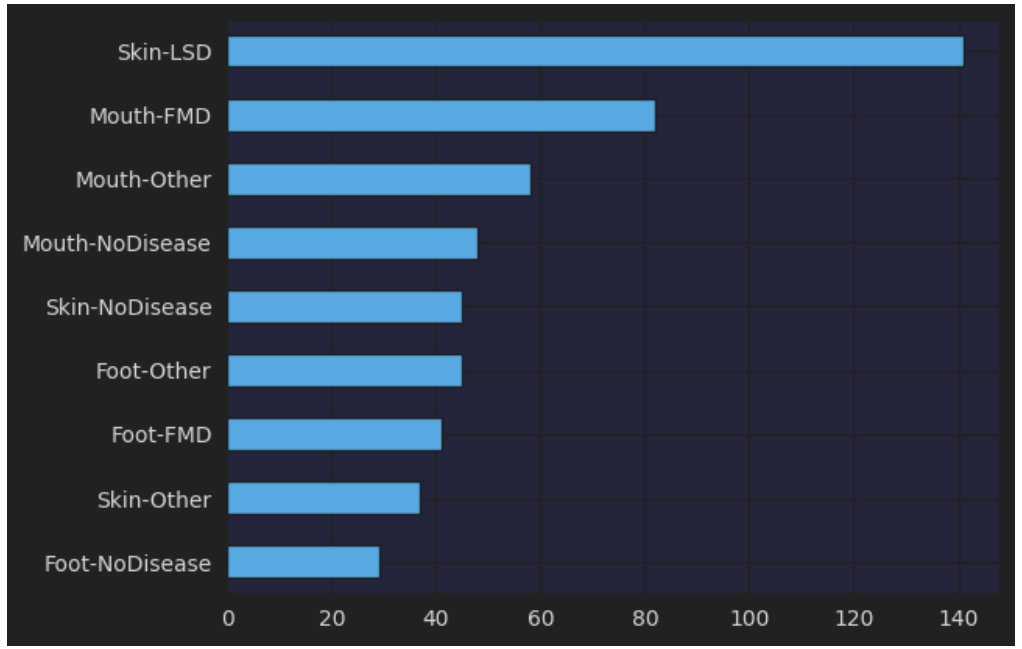
03.08.2023

PREVIOUS WORKS

- Model Development
- Web / Mobile Application Proof of Concept



4 Class Split

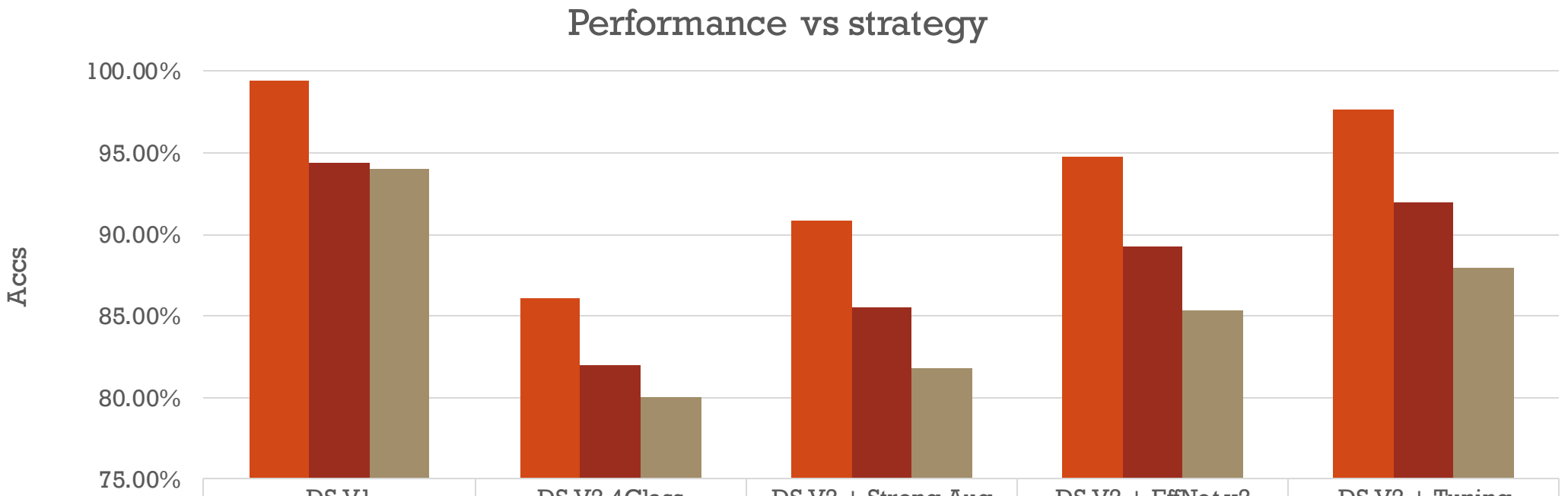


9 Class Split

NEW DATASET SPLIT POLICY

- We change the 4 classes classification problem to 9 classes. Leads to increased complexity and more accurate localization.

PERFORMANCE VS STRATEGY

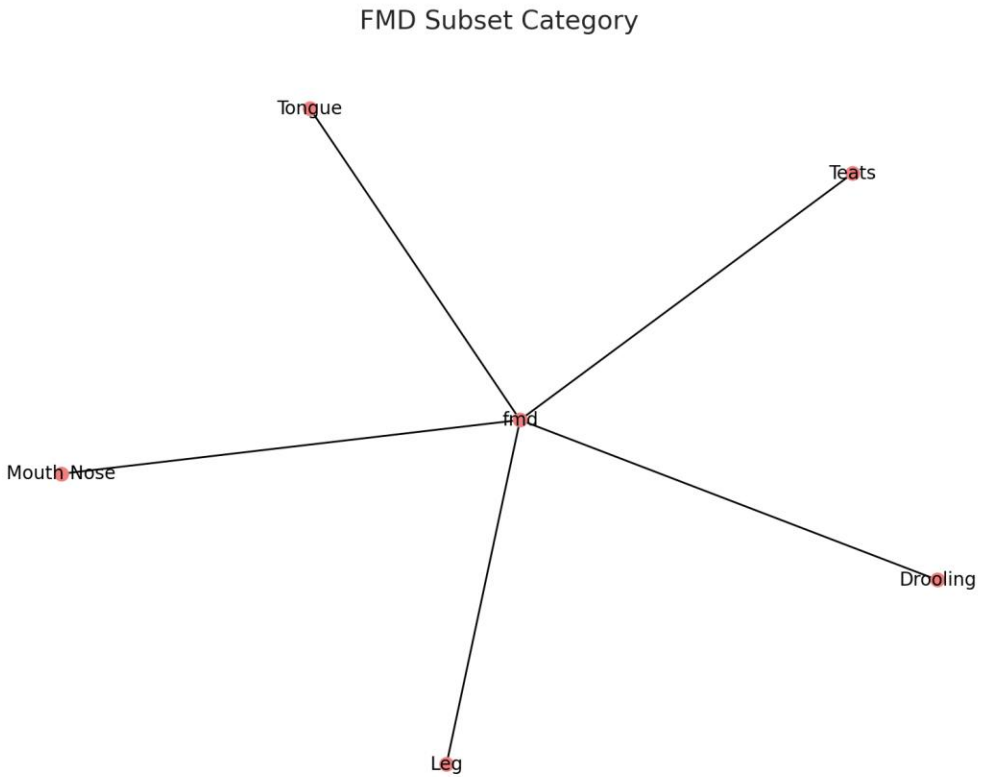


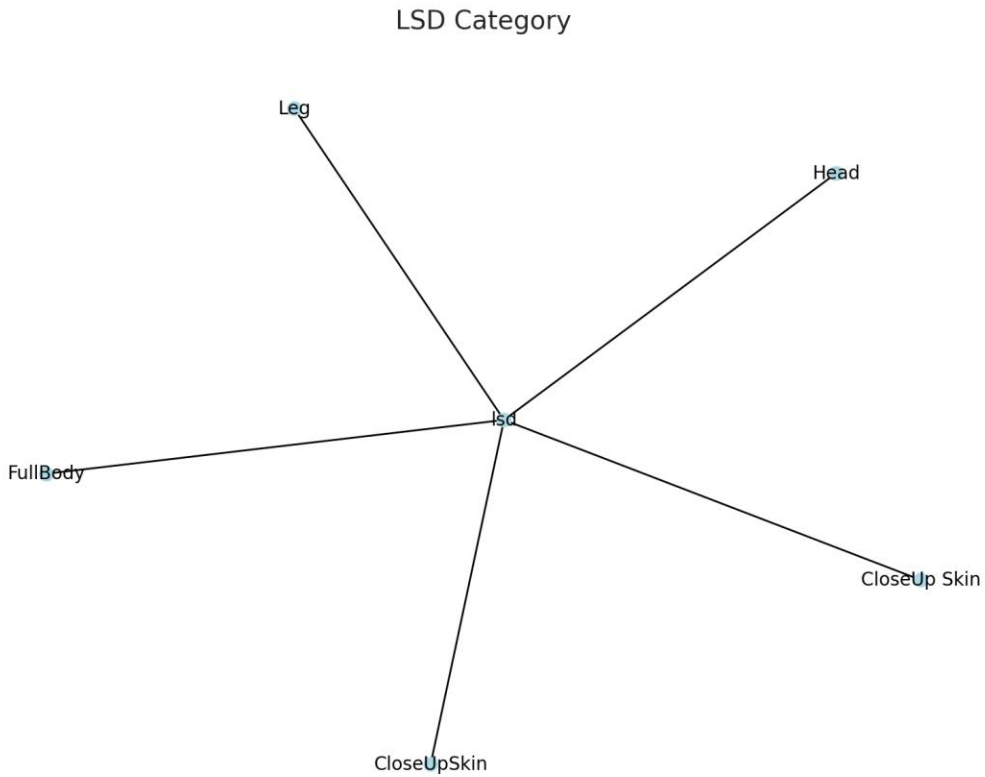
	DS V1	DS V2 4Class	DS V2 + Strong Aug	DS V2 + EffNet v2	DS V2 + Tuning
Train Acc	99.47%	86.13%	90.84%	94.75%	97.68%
Validation Acc	94.40%	81.99%	85.56%	89.24%	92.00%
Test Acc	94.00%	80.00%	81.83%	85.35%	87.99%

CONTINUOUS DEVELOPMENT

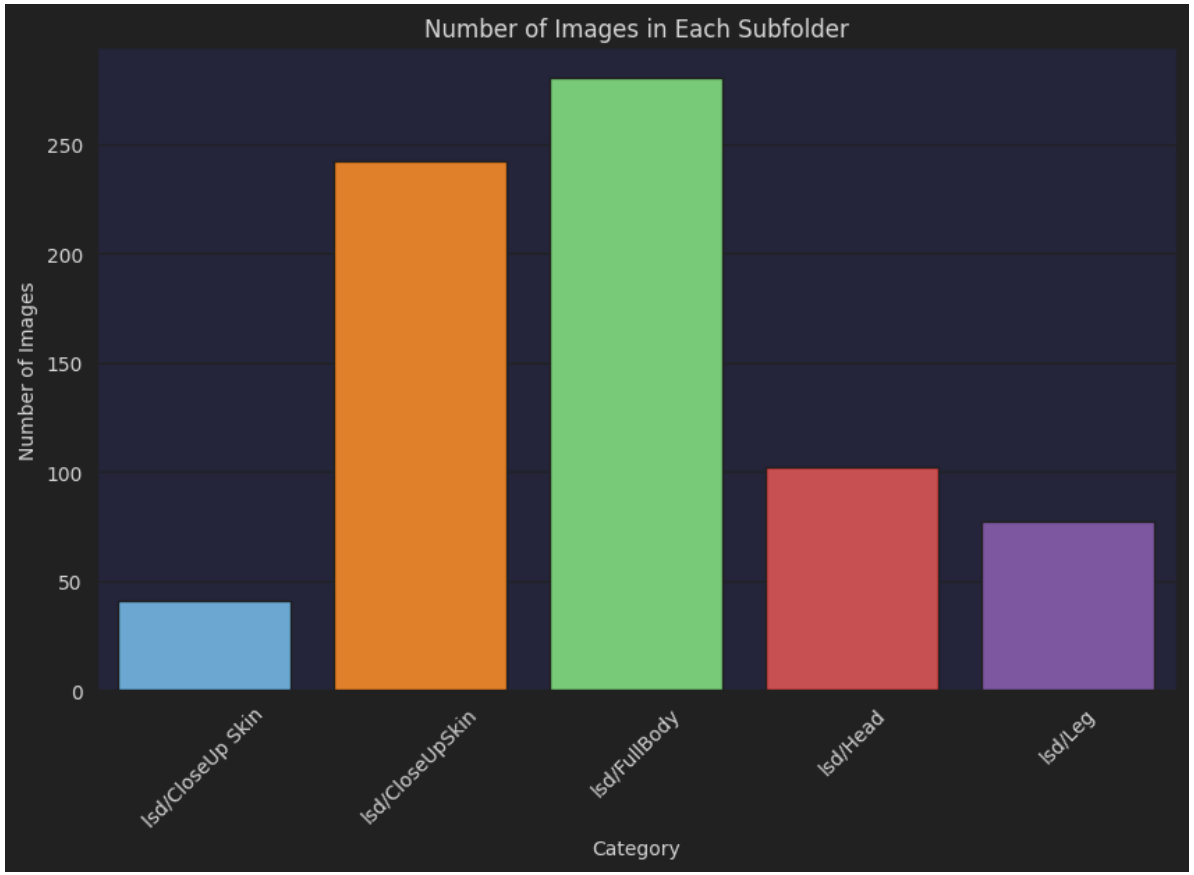
- Dataset Categories
- Failure Case analysis
- Context Understanding

DATASET CATEGORIES

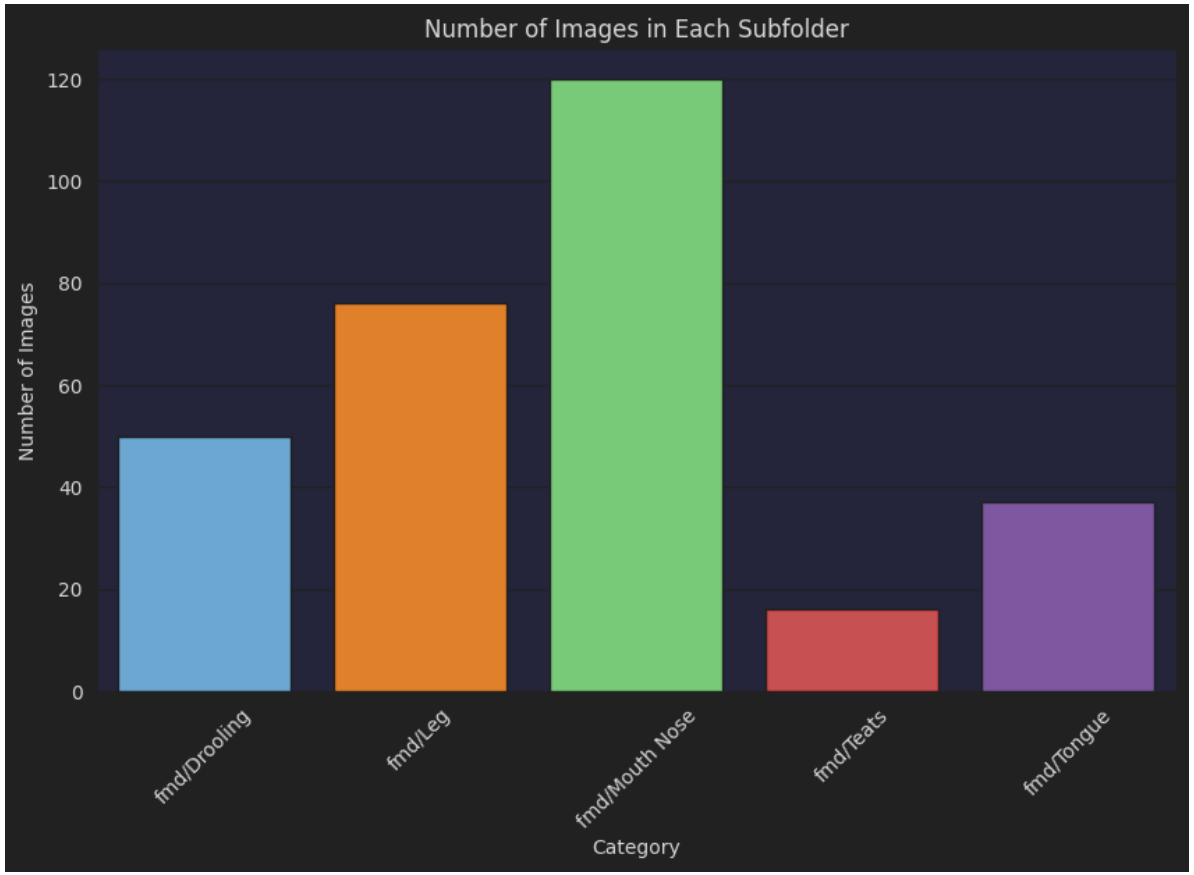




DATASET CATEGORIES



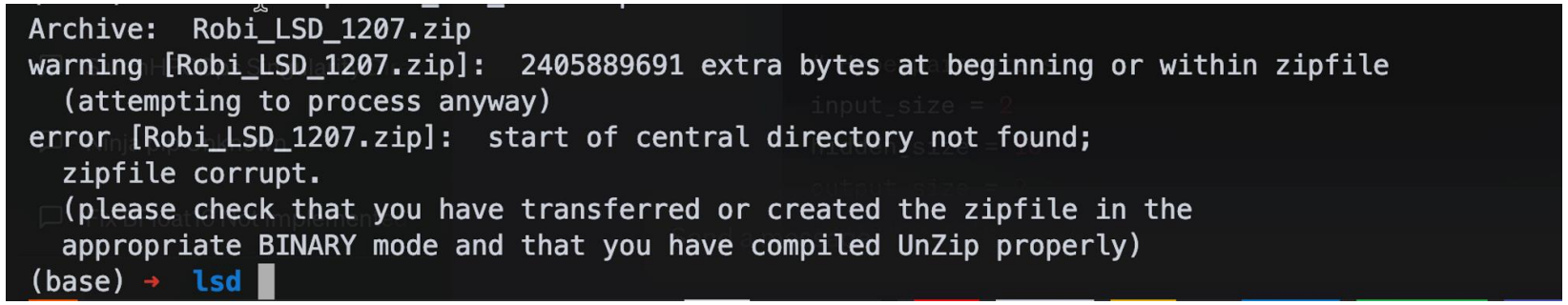
LSD



FMD

DATASET CATEGORIES

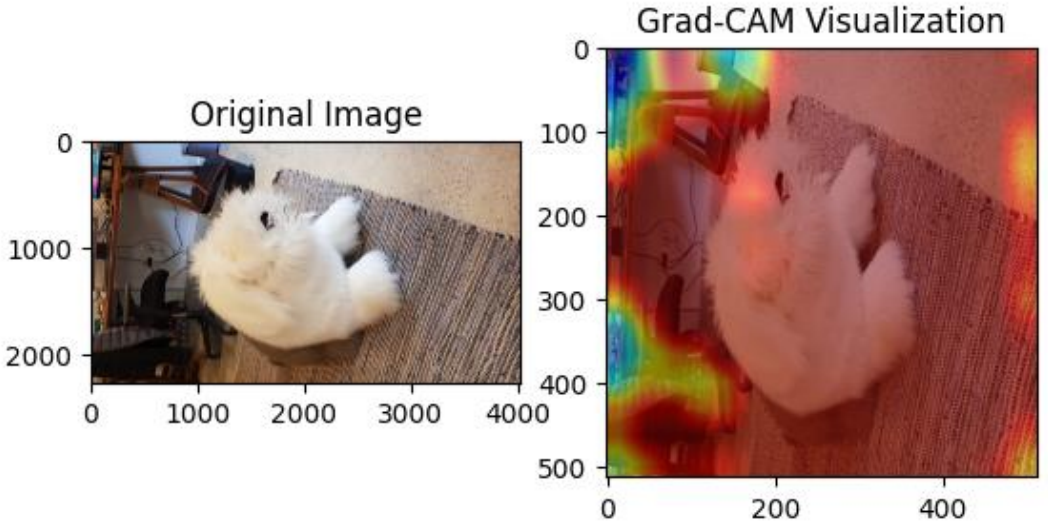
- Robi_LSD_1207.zip damaged



FAILURE CASE ANALYSIS CASE 1

- INPUT :

CAM (Attention Class FMD)

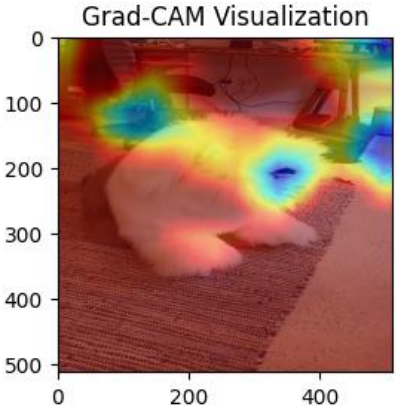


Prediction Score : [('Fmd', '70.820%'), ('No-Disease', '22.886%'), ('Lsd', '6.294%')]

CASE 1 INTERESTING FINDING

- Simple Solution by 270 degree rotation
- img = img.rotate(270, expand=True)

CAM (Attention Class FMD)

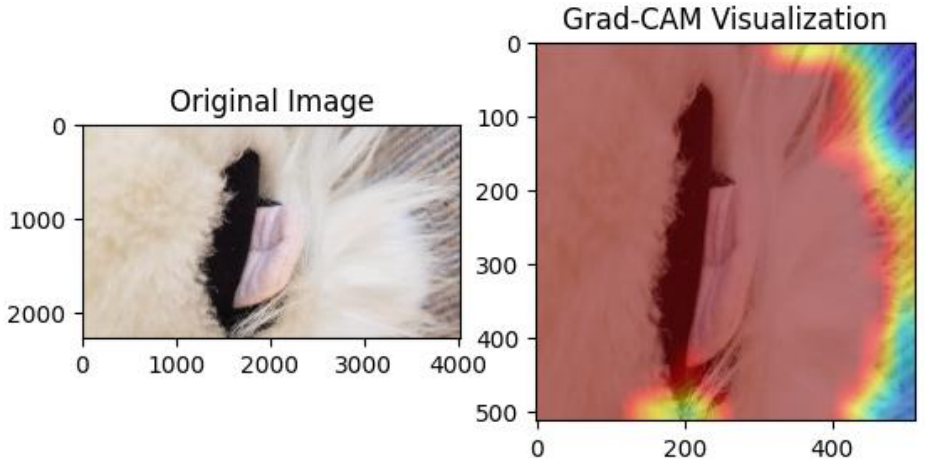


[('No-Disease', '72.605%'), ('Fmd', '19.244%'), ('Lsd', '8.151%')]

FAILURE CASE ANALYSIS CASE 2

- INPUT :

CAM (Attention Class FMD)

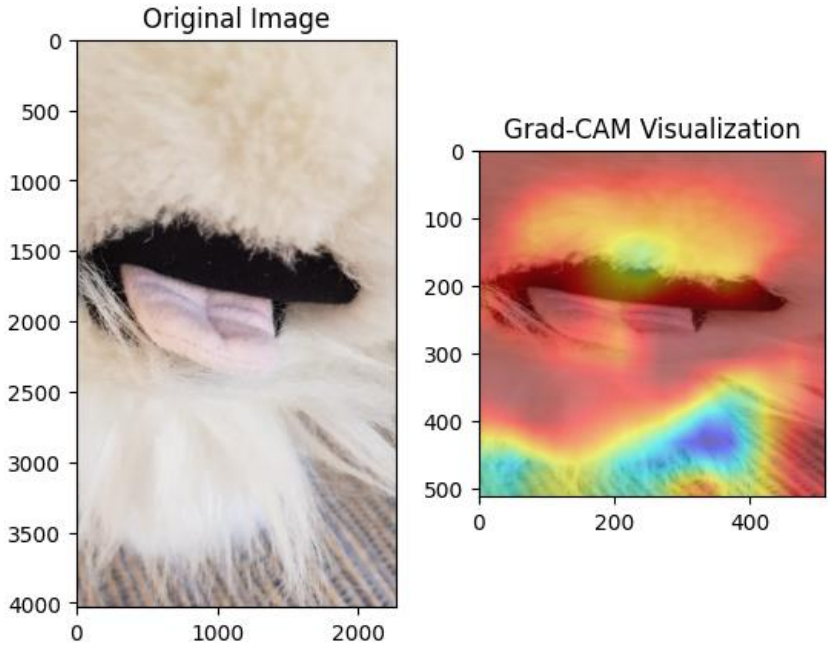


[('Fmd', '71.187%'), ('No-Disease', '28.616%'), ('Lsd', '0.197%')]

CASE 2 FIX

- Simple Solution by 270 degree rotation
- img = img.rotate(270, expand=True)

CAM (Attention Class FMD)

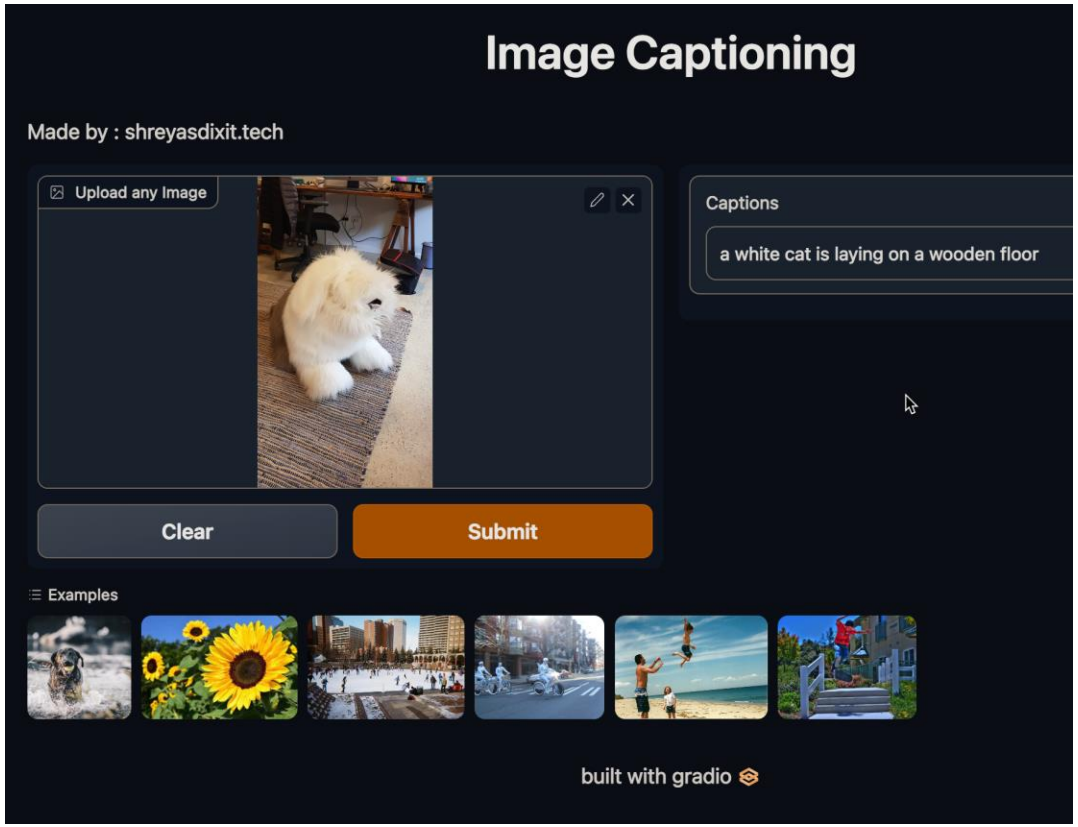


('No-Disease', '72.605%'),
('Fmd', '19.244%'),
('Lsd', '8.151%')

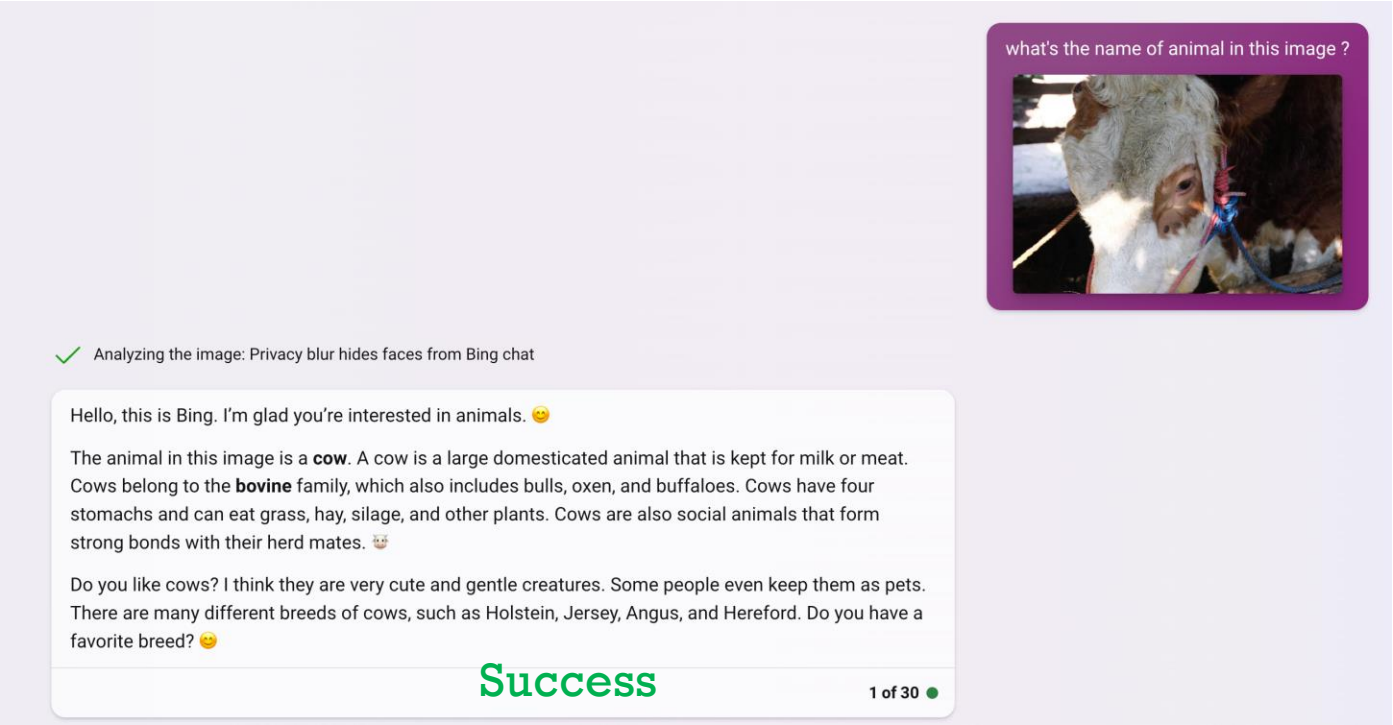
CONTEXT UNDERSTANDING

- Proposed Solution : Context Information Understanding **with Image Caption Model** or **Vision Awared LLMs**
- * This is *out* of current project score, we provide technical advice only.

IMAGE CAPTION MODEL

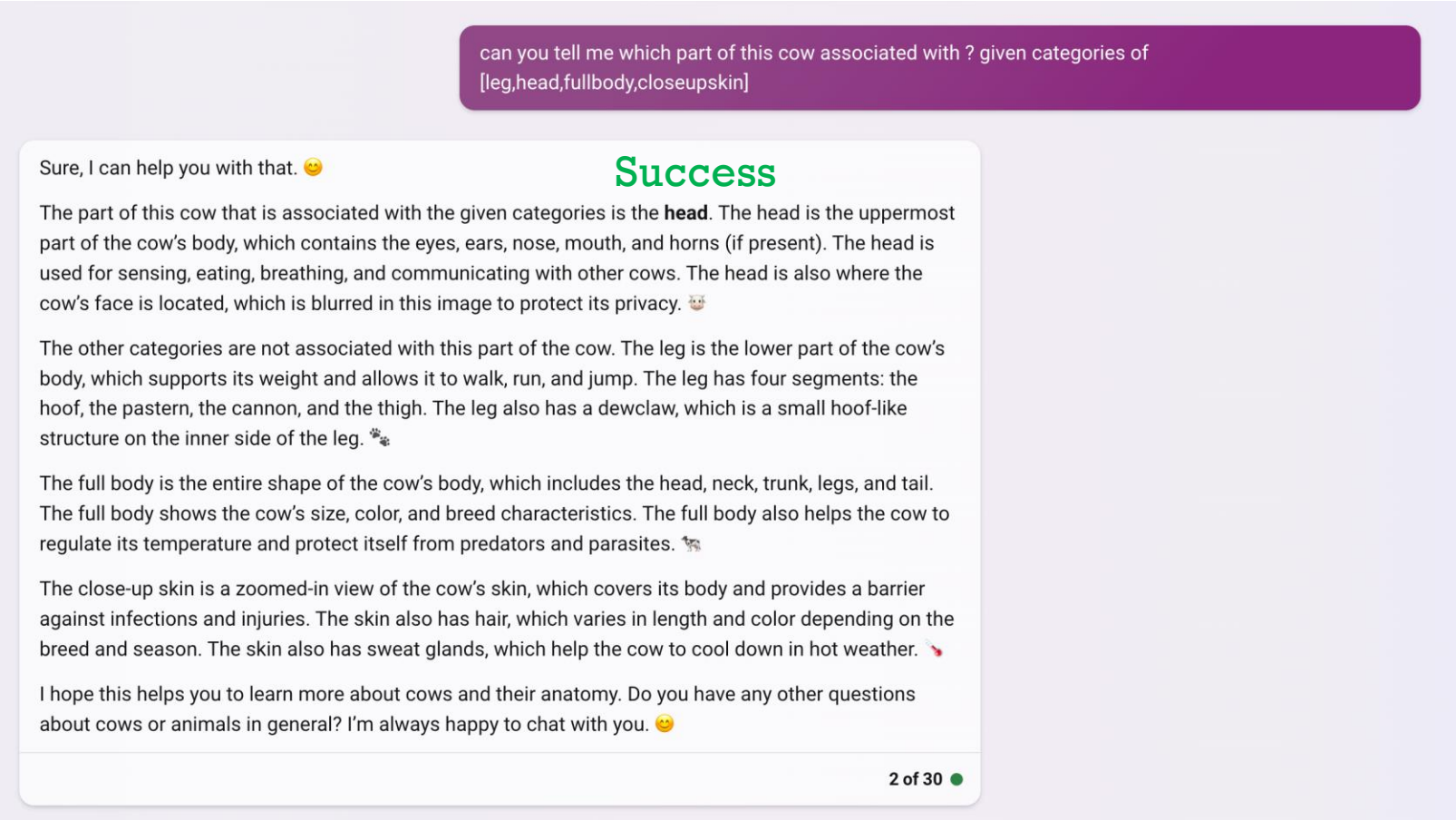


- Response : a white cat is laying on a wooden floor



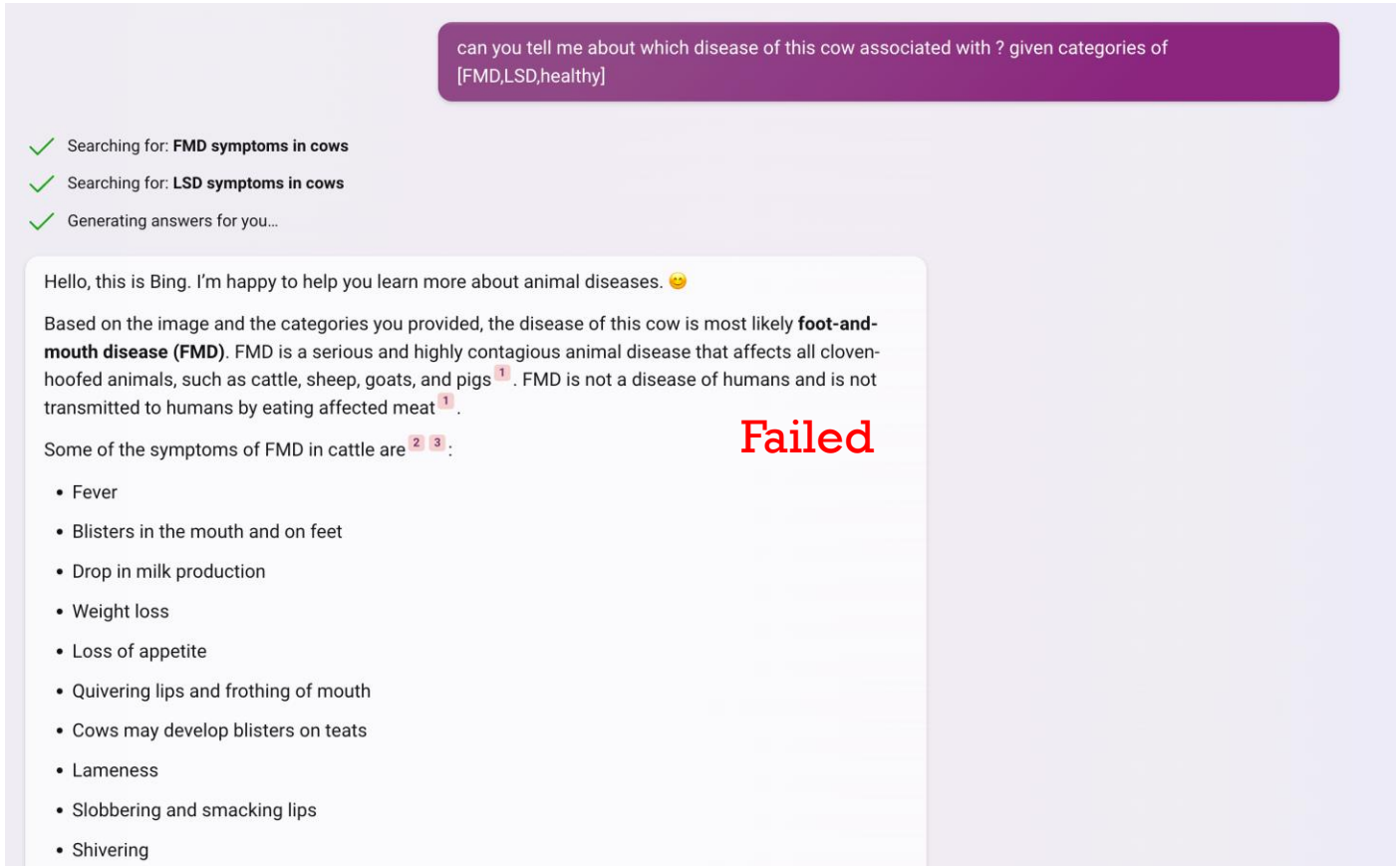
VISION AWARED LLMS

Context Understand 1 : Animal Category



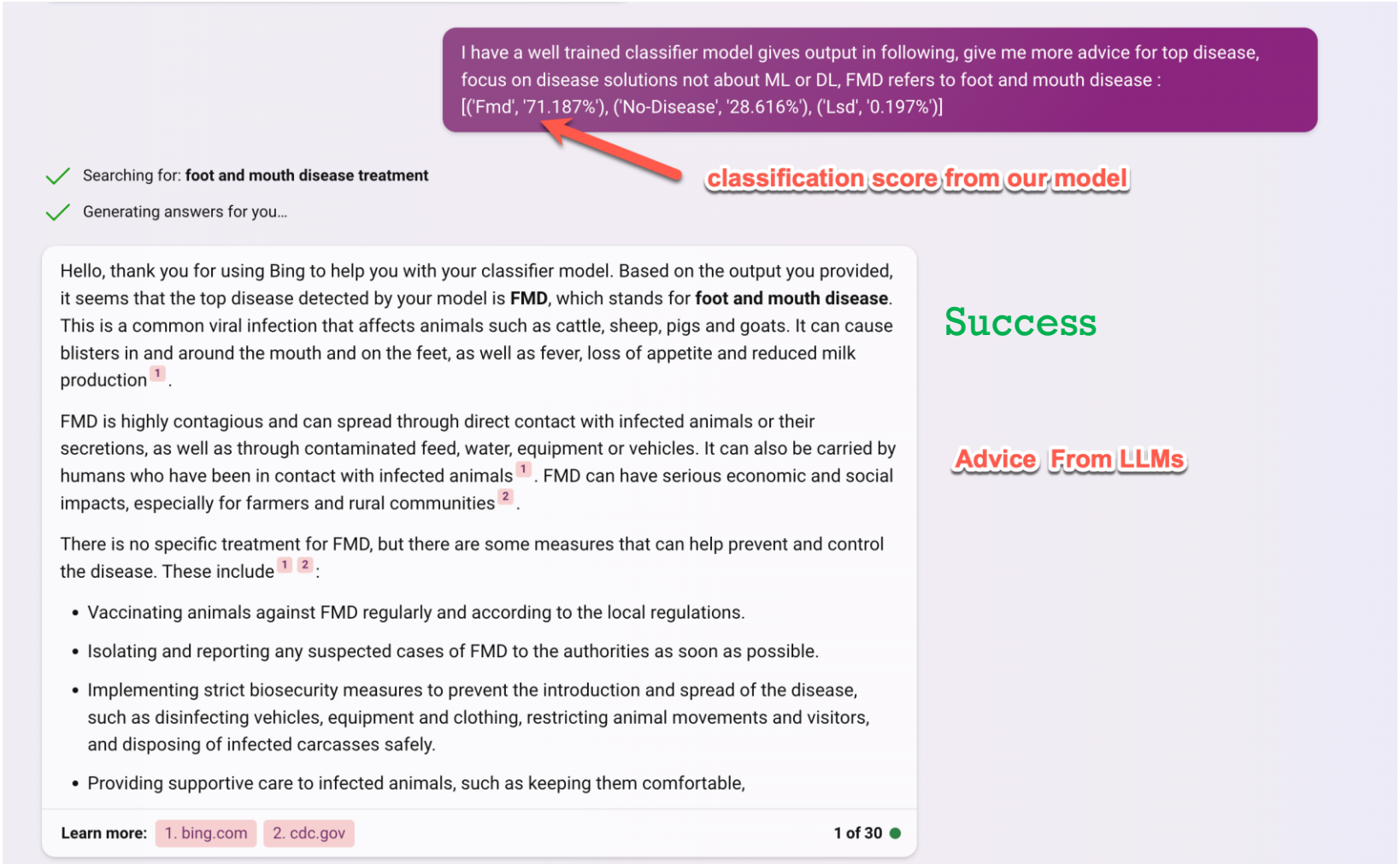
VISION AWARED LLMS

Context Understand 2 : Part Understanding



VISION AWARED LLMS

Context Understand 3 :
Disease Understanding

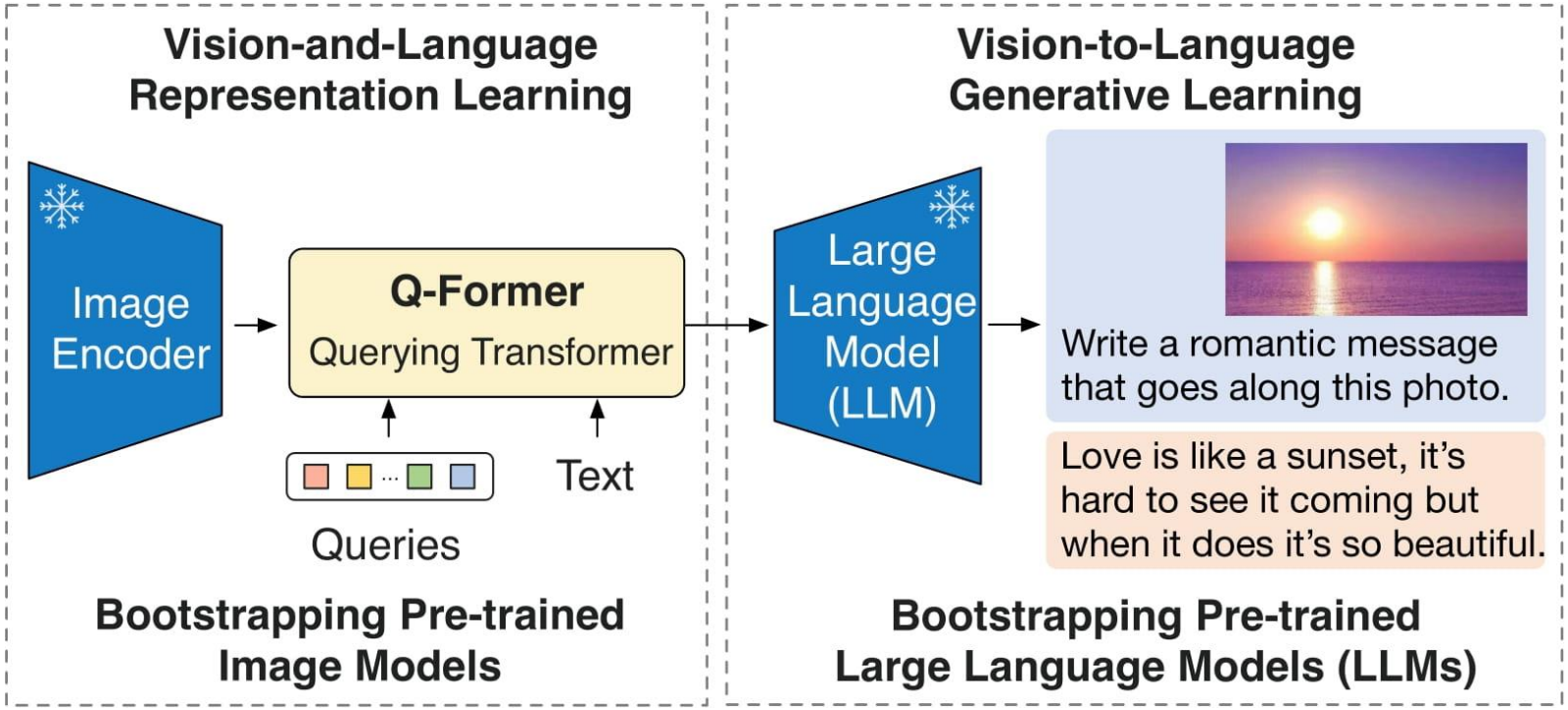


VISION AWARED LLMS

Context Understand 4 : Disease Understanding by external classifier

VISION AWARED LLMS : BLIP-2

- Open Source Alternative



Li, Junnan, et al. "Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models." arXiv preprint arXiv:2301.12597 (2023).

THANKS FOR WATCHING.

- Akide Liu
- Dr. Minh Hieu Phan
- Dr. YiFan Liu
- {akide.liu, vuminhhieu.phan,yifan.liu04}@adelaide.edu.au