OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

Developing a Deep Learning Classifier for Gout Management

Amir Bder
amir.bder.2009@gmail.com

ABSTRACT

Gout is a form of arthritis which results from hyperuricemia, the process where excess uric acid accumulates in the bloodstream. According to the NHANES survey, there are about 3.9% of total adults in the United States who do not suffer from gout, with 9.2 million individuals affected. In order for those suffering from gout to stay healthy, eating right and avoiding flare-ups will mean controlling one's diet and avoiding foods rich in purines. Unfortunately, dieting is quite troublesome. While eating meals, the majority of the individuals having gout have problems in looking at the plate of food and determining whether it contains dangerous amounts of purines.

As opposed to developing the usual ML model capable of classifying food into different categories, this work aims to develop a clinically relevant algorithm for categorizing food according to danger levels.

INTRODUCTION

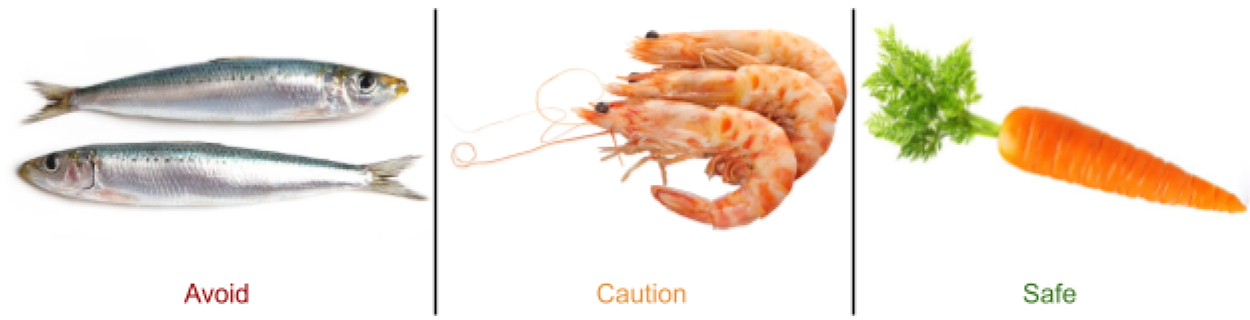
Gout is a form of arthritis which results from hyperuricemia (the process where excess uric acid accumulates in the bloodstream). According to the NHANES survey, there are about 3.9% of total adults in the United States who suffer from gout, with 9.2 million individuals affected. In order for those suffering from gout to stay healthy, eating right and avoiding flare-ups means controlling diet and avoiding foods rich in purines. Purines are chemical compounds found in certain foods that the body breaks down into uric acid. High levels of uric acid can crystallize in joints, causing the painful flare-ups associated with gout. Unfortunately, dieting is troublesome. While eating meals, the majority of the individuals having gout have problems in looking at the plate of food and determining whether it contains dangerous amounts of purines. Also, many high-purine ingredients are "hidden" or chemically embedded in complex meals (like meat-based stock in soup, or meat based gelatin in gummies), making visual identification challenging for patients to identify by themselves.

This study hypothesizes that a specialized binary deep learning classifier focusing on clinical risk (Avoid vs. Other) provides higher clinical utility for gout patients than standard multi-class food identification models. While apps like Yuka rely on scanning barcodes and labels for processed foods, this model can visualize prepared meals (like at restaurants) where labels are unavailable. As opposed to developing the usual Machine Learning (Machine Learning is a branch of technology that uses deep learning architectures to process and analyze data by imitating the way the human brain and eye identify visual patterns and features) model capable of classifying food into different categories. Standard ML Classification models identify categories of food items, such as “pizza”, “burger”, or “pasta” without health context. This work aims to develop a clinically relevant algorithm for categorizing food according to danger levels, translating a visual image into a "risk label" based on the Gout Knowledge Database (GKD).

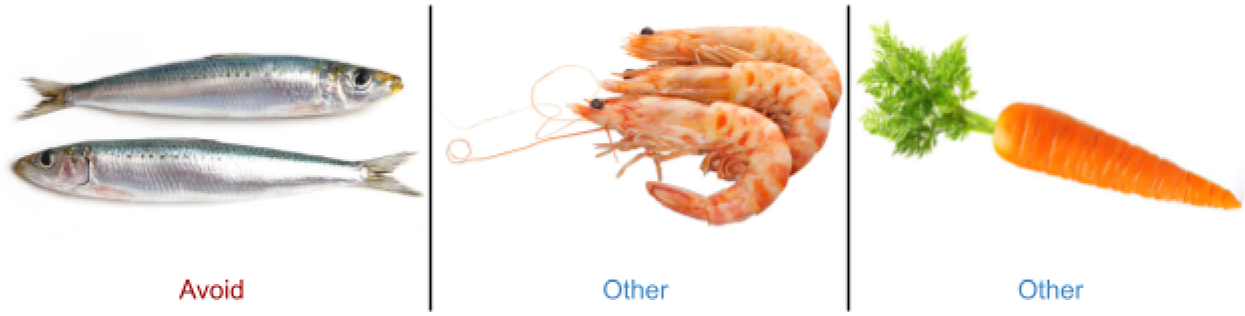
MATERIALS AND METHODS

This is based on the patient assistance-oriented system architecture. Initially, the user takes a picture of a particular meal through his/her mobile device. Then, the system utilizes deep learning to analyze the image based on its visual features and classify the meal. Afterward, the image will be checked against available chemical information concerning its purine content in grams. Based on this information, the deep learning algorithm makes decisions regarding whether the meal in question is a high-purine food that should be avoided by the patient or whether it belongs to any of the other categories. Eventually, this information will be relayed back to the patient.

The design was originally developed using a three-class system, which included labeling options of "Safe," "Caution," and "Avoid."



However, such an approach could not work effectively since there were no distinct features that could differentiate foods labeled "Caution" and "Safe" from each other. Thus, the design had to be changed to a two-class classification system of either "Avoid" or "Other" (the latter refers to "Safe" and "Caution").



In order to classify the food images into "Avoid" and "Non-Avoid" categories based on the visual features, a Convolutional Neural Network (CNN), which comprises Conv2D and MaxPooling layers, is utilized. Although this term can be too complex, it refers to a special deep learning architecture designed specifically for processing visual data and imitating the work of the human eye. In contrast to regular neural networks, where images are treated as matrices of numbers, a CNN maintains the relationships between pixels by utilizing "filters". These filters work as scanners scanning the input data in order to discover its visual features including edges, textures, and shapes. It's important to note that the algorithm doesn't directly see the purine level molecularly, it identifies the textures and patterns identifiable in a certain food item, such as sections, opacity, or other features, associating it with its respective purine level. By applying pooling operation, the model reduces dimensionality of the data during further transformation. In case of MaxPooling, a matrix scans the input and picks the highest value from each area. Thus, only the most salient visual features are retained while the extra background is discarded. Finally, the extracted features are flattened and sent to the classification output layer which computes the probability of belonging to the "Avoid" class according to the visual texture.

To enhance the performance, we created and applied training methods based on dynamic data augmentation. Randomly flipping and rotating images teaches the neural net to recognize the dish irrespective of its position. Dynamic data augmentation means that the adjustment occurs dynamically throughout the process of training, as opposed to adjusting test/train split. In order to combat overfitting and prevent the network from memorizing training images, a Dropout layer with a dropout rate p=0.5 was added. The selection of p=0.5 is typical and represents a great baseline because it introduces the highest amount of noise by randomly setting 50% of neuron activations to zero in each step of the training process.

DATASET AND KNOWLEDGE MAPPING

The basis of the current study can be characterized by the creation of an innovative dataset tailored for the specific purpose of analyzing the needs of gout patients. Instead of relying on the general dataset available in the Food-101 archive, a specific purine-indexed dataset has been created to cater to the needs of clinical applications. Whereas the general-purpose Food-101 dataset does not have any clinical context, the most important original contribution that can be made in this phase was the creation of the Gout Knowledge Database (GKD).

G - Gout-Relevant (54 Items)

These items are generally high or moderate in purines and are known triggers for uric acid spikes.

- **baby_back_ribs** - G (pork ribs = moderate-high purine, RELEVANT)
- **beef_carpaccio** - G (raw beef = high purine, RELEVANT)
- **beef_tartare** - G (raw beef = high purine, RELEVANT)
- **bibimbap** - G (typically contains beef/egg, RELEVANT)
- **breakfast_burrito** - G (usually contains sausage/bacon, RELEVANT)
- **caesar_salad** - G (anchovies in dressing = very high purine, RELEVANT)
- **ceviche** - G (seafood/shellfish = high purine, RELEVANT)
- **chicken_curry** - G (chicken = moderate purine, RELEVANT)
- **chicken_quesadilla** - G (chicken = moderate purine, RELEVANT)
- **chicken_wings** - G (chicken/skin = moderate-high purine, RELEVANT)
- **clam_chowder** - G (clams/shellfish = high purine, RELEVANT)
- **club_sandwich** - G (bacon/turkey/ham = moderate-high purine, RELEVANT)
- **crab_cakes** - G (crab/shellfish = high purine, RELEVANT)
- **croque_madame** - G (ham = moderate-high purine, RELEVANT)
- **dumplings** - G (usually pork/shrimp filling = high purine, RELEVANT)
- **eggs_benedict** - G (canadian bacon/ham = moderate-high purine, RELEVANT)
- **escargots** - G (mollusks = moderate-high purine, RELEVANT)
- **filet_mignon** - G (beef = high purine, RELEVANT)
- **fish_and_chips** - G (white fish/frying = moderate-high purine, RELEVANT)
- **foie_gras** - G (organ meat = extremely high purine, RELEVANT)
- **french_onion_soup** - G (beef stock/broth = moderate-high purine, RELEVANT)
- **fried_calamari** - G (squid/shellfish = high purine, RELEVANT)
- **fried_rice** - G (usually contains pork/shrimp, RELEVANT)
- **grilled_salmon** - G (oily fish = moderate-high purine, RELEVANT)
- **gyoza** - G (pork/shrimp filling = high purine, RELEVANT)
- **hamburger** - G (beef = high purine, RELEVANT)
- **hot_and_sour_soup** - G (usually pork/shrimp/broth, RELEVANT)
- **hot_dog** - G (processed meat = high purine, RELEVANT)
- **lasagna** - G (beef/sausage = moderate-high purine, RELEVANT)
- **lobster_bisque** - G (shellfish/stock = high purine, RELEVANT)
- **lobster_roll_sandwich** - G (lobster/shellfish = high purine, RELEVANT)
- **miso_soup** - G (dashi/fish stock base = moderate-high purine, RELEVANT)
- **mussels** - G (shellfish = very high purine, RELEVANT)
- **nachos** - G (usually topped with beef/chili, RELEVANT)
- **oysters** - G (shellfish = very high purine, RELEVANT)
- **pad_thai** - G (shrimp/fish sauce = high purine, RELEVANT)
- **paella** - G (mixed seafood/sausage = high purine, RELEVANT)
- **peking_duck** - G (duck = moderate-high purine, RELEVANT)
- **pho** - G (beef/bone broth = high purine, RELEVANT)
- **pizza** - G (if topped with meat like pepperoni/sausage, RELEVANT)
- **pork_chop** - G (pork = high purine, RELEVANT)
- **poutine** - G (beef gravy = high purine, RELEVANT)
- **prime_rib** - G (beef = high purine, RELEVANT)
- **pulled_pork_sandwich** - G (pork = high purine, RELEVANT)
- **ramen** - G (pork/bone broth = high purine, RELEVANT)
- **ravioli** - G (if meat-filled = moderate-high purine, RELEVANT)
- **risotto** - G (usually contains chicken/beef stock, RELEVANT)
- **samosa** - G (usually contains meat/moderate purine, RELEVANT)
- **sashimi** - G (raw fish = high purine, RELEVANT)
- **scallops** - G (shellfish = very high purine, RELEVANT)
- **shrimp_and_grits** - G (shrimp = high purine, RELEVANT)
- **spaghetti_bolognese** - G (meat sauce = high purine, RELEVANT)
- **spaghetti_carbonara** - G (pancetta/pork = high purine, RELEVANT)
- **spring_rolls** - G (usually contain shrimp/pork, RELEVANT)
- **steak** - G (beef = high purine, RELEVANT)
- **sushi** - G (fish = high purine, RELEVANT)
- **tacos** - G (usually beef/pork/shrimp, RELEVANT)
- **takoyaki** - G (octopus = high purine, RELEVANT)
- **tuna_tartare** - G (tuna = high purine, RELEVANT)

N - Not Gout-Relevant (47 Items)

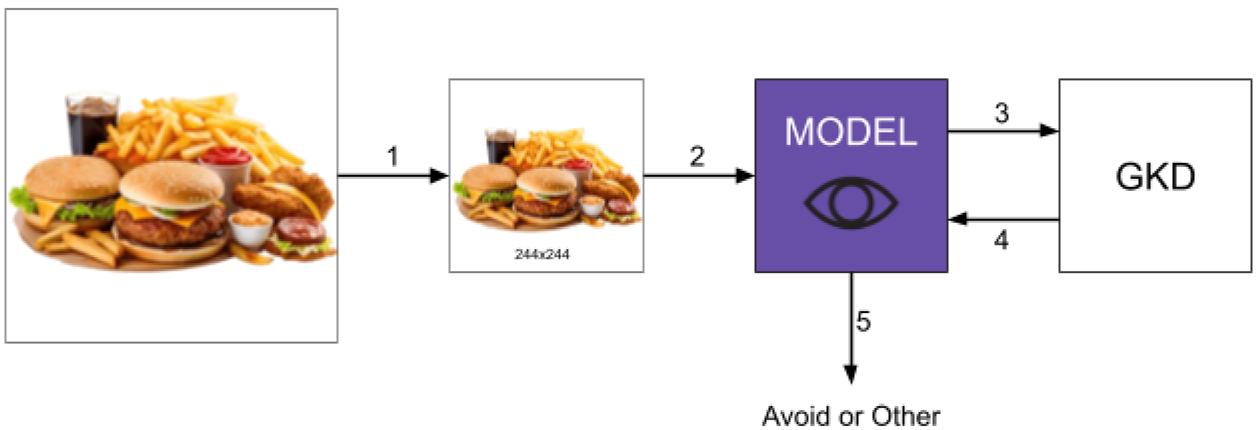
These items are low in purines and generally considered safe for those managing gout.

- **apple_pie** - N (dessert, not gout-relevant)
- **baklava** - N (pastry, low purine)
- **beet_salad** - N (vegetable, low purine)
- **beignets** - N (fried dough, low purine)
- **bread_pudding** - N (dairy/bread, low purine)
- **bruschetta** - N (tomatoes/bread, low purine)
- **cannoli** - N (dairy pastry, low purine)
- **caprese_salad** - N (mozzarella/tomato, low purine)
- **carrot_cake** - N (vegetable/cake, low purine)
- **cheesecake** - N (dairy, low purine)
- **cheese_plate** - N (dairy is low-purine, not gout-relevant)
- **chocolate_cake** - N (dessert, low purine)
- **chocolate_mousse** - N (dessert, low purine)
- **churros** - N (fried dough, low purine)
- **creme_brulee** - N (custard, low purine)
- **cup_cakes** - N (dessert, low purine)
- **deviled_eggs** - N (eggs are low purine, not gout-relevant)
- **donuts** - N (fried dough, low purine)
- **edamame** - N (soy/legumes are low-moderate but generally safe)
- **falafel** - N (chickpeas, low purine)
- **french_fries** - N (potatoes, low purine)
- **french_toast** - N (bread/egg, low purine)
- **frozen_yogurt** - N (dairy, low purine)
- **garlic_bread** - N (bread, low purine)
- **gnocchi** - N (potato pasta, low purine)
- **greek_salad** - N (vegetables/feta, low purine)
- **grilled_cheese_sandwich** - N (dairy/bread, low purine)
- **guacamole** - N (avocado, low purine)
- **huevos_rancheros** - N (beans/eggs/tortilla, low purine)
- **hummus** - N (chickpeas, low purine)
- **ice_cream** - N (dairy, low purine)
- **macaroni_and_cheese** - N (dairy/pasta, low purine)
- **macarons** - N (almond/egg white, low purine)
- **omelette** - N (eggs, low purine)
- **onion_rings** - N (vegetable, low purine)
- **pancakes** - N (batter, low purine)
- **panna_cotta** - N (cream, low purine)
- **red_velvet_cake** - N (dessert, low purine)
- **seaweed_salad** - N (seaweed, low purine)
- **strawberry_shortcake** - N (fruit/cake, low purine)
- **tiramisu** - N (dessert, low purine)
- **waffles** - N (batter, low purine)

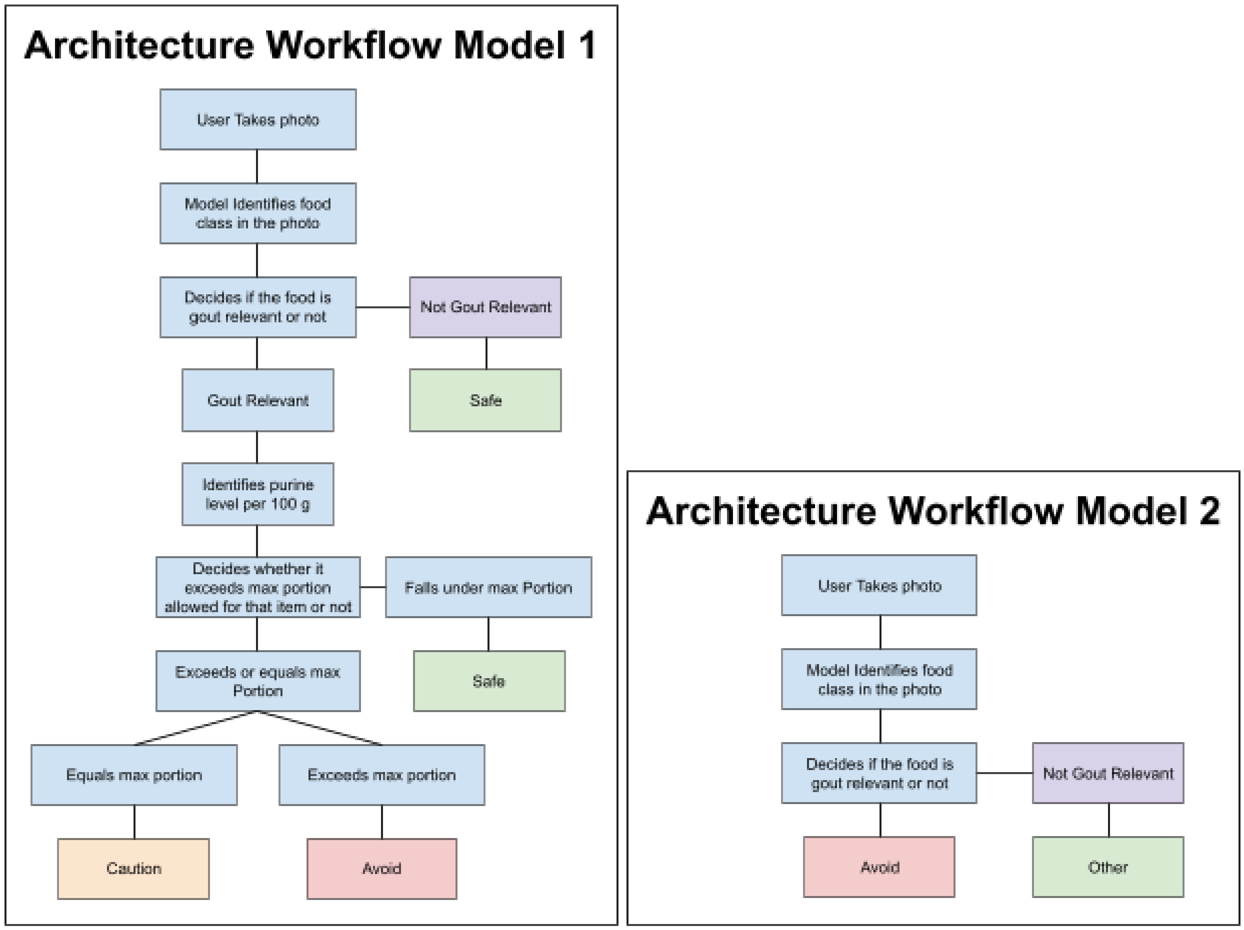
The system utilizes a two step mapping process. First, it identifies the food category, then it cross references it with the GKD I created, which contains pre-indexed data separating between gout relevant (Avoid) and non gout relevant (Other). This allows the system to provide risk assessment based on established data rather than trying to chemically analyze an image.

The creation of the GKD took place in three stages. First, through a process of selective filtration, 54 specific food categories were handpicked from the original 101 food types because of the frequency of appearance of these food items in the clinical literature on gout. Second, a process of purine indexing was undertaken wherein each of the filtered 54 food categories underwent an analysis that measured purine density in milligrams per 100 grams of food in a laboratory setting. Third, the images underwent a process of clinical labeling to allow transition from "food naming" to "risk labeling".

For the model to accommodate for the complexities of different food textures while keeping the computational requirements low, a uniform preprocessing process was designed. All input images were subjected to spatial normalization to make them have a resolution of 224x224 pixels.



This particular size was selected to allow for sufficient information regarding the distinction between the two categories while at the same time limiting the computational cost on the training environment. The next step involved feature scaling that changed pixel values from the initial interval of [0;255] to [0;1]. Normalization of the inputs allows for the distribution to be uniform, thus avoiding gradient explosions and substantially speeding up the training phase of the CNN. Last but not least, class balancing strictly maintained the balance between the two classes ("Avoid" and "Others").



RESULTS AND DISCUSSION

The test results prove that simplification itself leads to better practical outcomes. The considerable improvement in all the figures observed during the switch to the binary approach illustrates the notion that an AI system functions best when it concentrates on the key moment of the human decision.

By introducing the weighting according to clinical importance into the cost function, we showed the model that failing to detect a high-purine foodstuff is much more serious than being overly cautious about a non-dangerous one. Statistical overview of this achievement is provided in the table below:

Multi-Class vs. Binary Model Performance Metrics		
Metric	**3-Class Model (Prototype)**	**2-Class Model (Current)**
Accuracy	0.39	0.56
Precision	0.34	0.54
Recall (Sensitivity)	0.31	0.58
F1-Score	0.33	0.56
AUC-ROC	0.51	0.59

Considering these figures, the key breakthrough for a clinical perspective here lies in raising the Recall to 0.58. The sensitivity, which indicates that a higher percentage of potentially hazardous triggers are now being identified by the classifier, is crucial in the medical domain due to class weighting penalizing false negatives. The significant increase in overall accuracy from 39% to 56% justifies the choice of moving away from the multi-class format.

However, it should also be mentioned that there is a critical downside to these achievements, since dietary health risk is extremely difficult to enhance based on purely visual data. There is an invisible limitation in the capacity of computer vision algorithms, since many purines are not visible but rather embedded chemically within the image of food. This means that cream soup made using vegetable base would look virtually the same as a cream soup made using a highly purine meat/seafood stock. As the chemical components of such purine are impossible to distinguish solely from visual appearance, this sets a limitation on how accurate a classification algorithm may become. Future steps include trying Transfer Learning with ResNet/MobileNetV2.

LIMITATIONS

A big limitation to this approach is its inability to distinguish between ingredient quality. Itcan't extinguish between fresh meals, as opposed to a meal made with processed components, as chemical materials in processed food can impact gout differently. Because the model relies on visual identification of the final dish, it cannot yet analyze raw ingredients or text-based menu descriptions before a meal is prepared. As well as that, the current model isn't accurate enough to be plausibly used in real time. As of current testing, the accuracy yield is not yet high enough to concur that the model can detect foods less visible in the image, and occasionally struggles to detect a multitude of food products in an image. Also, if the food item isn't currently in the database I created, it won't take ingredient's lists into consideration.

NEXT STEPS & FUTURE DIRECTIONS

The objective for improving the clinical utility of the system is through the use of more sophisticated deep learning approaches that go beyond the current custom scratch-built CNNs. There are four key aspects:

- **Transfer Learning:** The future iterations of the app will make use of sophisticated ResNet50 or MobileNetV2 neural networks pre-trained on the gargantuan ImageNet database. Leveraging "pre-trained visual brains" means not wasting computing power on simple shapes recognition and concentrating on recognizing complex textures of the foods involved.
- **Phones:** Given the high performance to size ratio of MobileNetV2, preference is given to using this particular architecture, which will enable running the classification tool on the local device, i.e., a patient's phone. That will mean immediate classification results, regardless of available internet speed.
- **Fine-Tuning:** Selective unfreezing of the upper layers will be used to fine-tune weights for maximum specificity, enabling differentiation between visually similar yet chemically different foods, for example, between high-purine animal organs and more benign proteins.
- **Integration & Metadata:** As mentioned earlier, to solve the problem of dependence on solely computer vision, the system will integrate an ensemble metadata analysis to add crucial information about what the patient might be eating or thinking of eating.
- **Gout Knowledge Database (GKD) Expanding:** Current efforts will focus on growing our custom training set by sampling larger sets of key "avoid" groups and adding much more varied cuisines from all around the world.

CONCLUSION

This study demonstrates the feasibility of moving towards specialization within computer vision models. Though the current accuracy rate of 55.74% highlights how difficult it is to distinguish between similar foods, the drastic improvements seen between the prototype and subsequent versions, with a tri-label design showing an impressive leap in accuracy, prove the wisdom behind choosing a binary classification focused on risk. By concentrating the model's efforts on the most essential human decision, that is Avoid versus other, and through careful application of clinical weighting techniques, this study provides an effective model prioritizing patient safety over accuracy alone. Results indicate that although a unique CNN can be used for detecting basic geometrical shapes of food products, the intricacies of clinical dietary assessment need the sophistication of pre-trained architectures. In conclusion, the classifier created here is a good proof-of-concept for the mobile application aimed at assisting gout patients.

* It is important to note that the model I created is designed specifically to detect only food items in images, not words or sets of lists.

REFERENCES

Aihemaitijiang, Sideer, et al. "The Association between Purine-Rich Food Intake and Hyperuricemia: A Cross-Sectional Study in Chinese Adult Residents." *Nutrients*, vol. 12, no. 12, Dec. 2020, p. 3835, https://doi.org/10.3390/nu12123835.

Butler, et al. *PubMed / MDPI Article on Epidemiology & NHANES Gout Statistics*. 2021.

Qaseem, Amir, et al. "Management of Acute and Recurrent Gout: A Clinical Practice Guideline From the American College of Physicians." *Annals of Internal Medicine*, vol. 166, no. 1, Jan. 2017, pp. 58–68, https://doi.org/10.7326/M16-0570.

"A Novel Method for Accurate Real-Time Food Classification: The Synergistic Integration of EfficientNetB7, CBAM, Transfer Learning and Data Augmentation." *ResearchGate*, Oct. 2024, https://www.researchgate.net/publication/384630532_A_Novel_Method_for_Accurate_Real-time_Food_Classification_The_Synergistic_Integration_of_EfficientNetB7_CBAM_Transfer_Learning_and_Data_Augmentation.