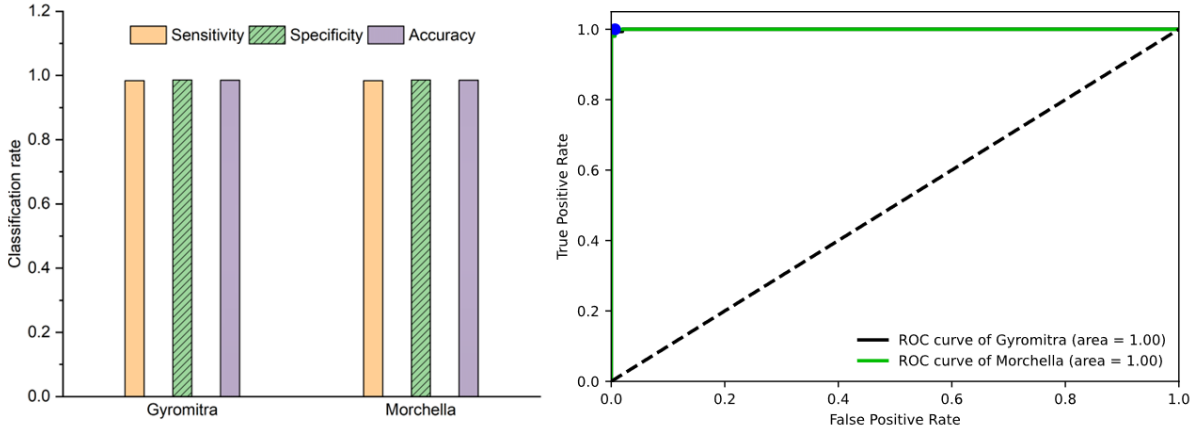
Machine learning-based classification of mushrooms using a smartphone application

Jae Joong Lee[1], Catherine Aime[2], Bartek Rajwa[3], and Euiwon Bae[4]*

[1]School of Computer Science, [2]Department of Plant and Botany Science, [3]Bindley Bioscience Center, [4]School of Mechanical Engineering, Purdue University, West Lafayette, Indiana 47907, USA

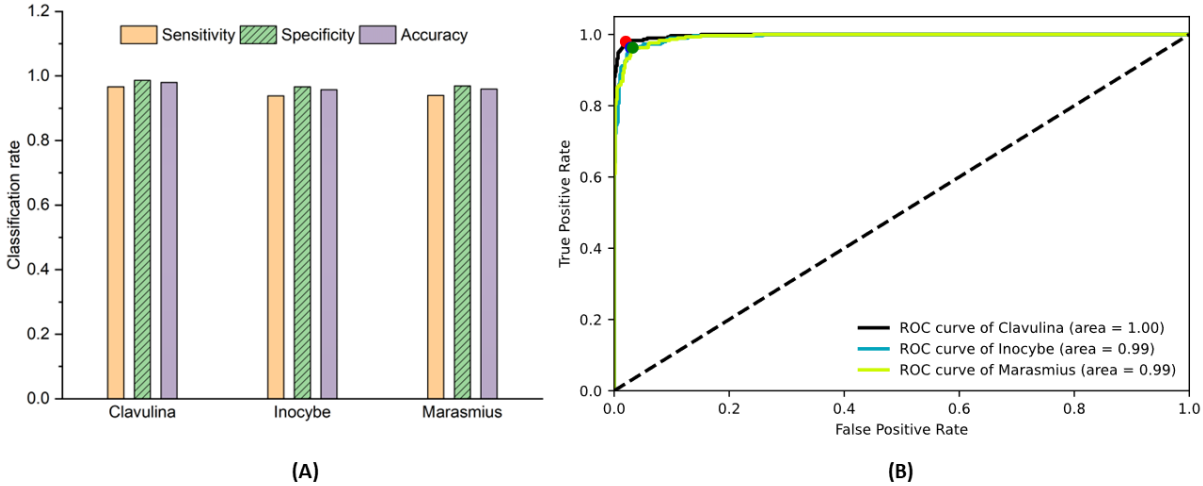
**Corresponding author: ebae@purdue.edu*

Supplementary material



(A) **(B)**

Figure S1. Comparison of sensitivity and specificity for the two classes of *Gyromitra* and *Morchella*. **(A)** classification rate from server-based training and **(B)** receiver operating curve.



(A) **(B)**

Figure S2. Comparison of sensitivity and specificity for the three classes of *Clavulina*, *Inocybe*, and *Marasmius*. **(A)** classification rate from server-based training and **(B)** receiver operating curve.

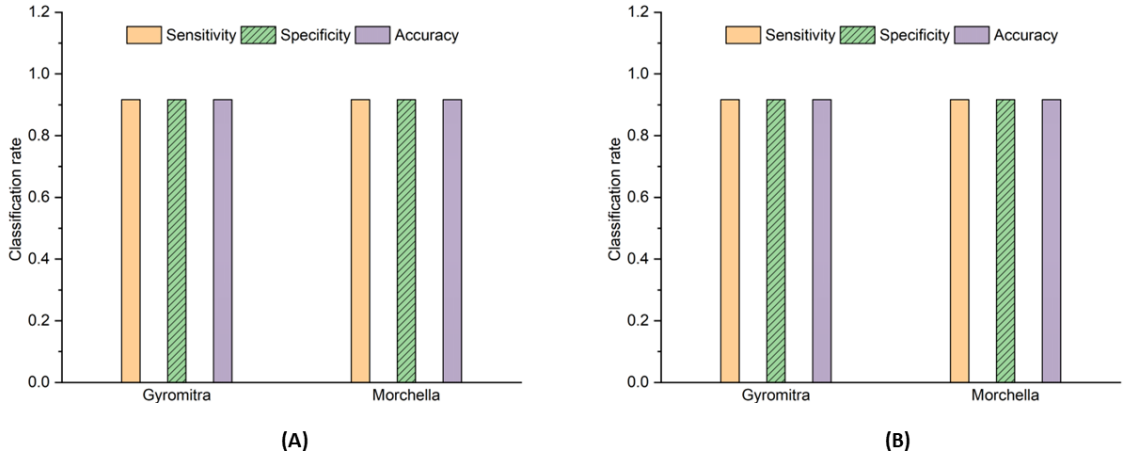


Figure S3. Comparison of sensitivity and specificity for the two classes of *Gyromitra* and *Morchella*. **(A)** classification rate from laptop training and testing. **(B)** test result from importing the trained model to the smartphone handset.

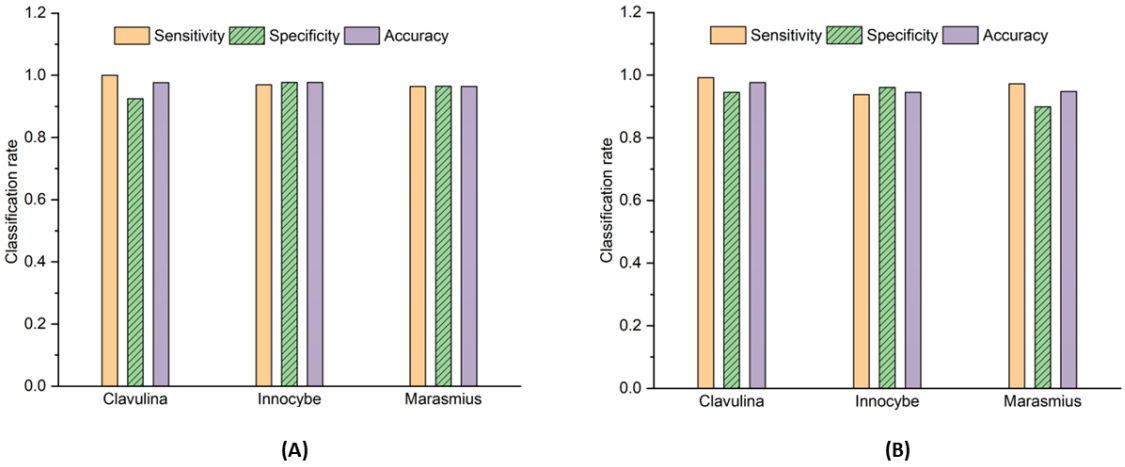


Figure S4. Comparison of sensitivity and specificity for *Clauvulina*, *Inocybe*, and *Marasmius* classes. **(A)** classification rate from laptop training and testing. **(B)** test result from importing the trained model to the smartphone handset.

Table S1. Computed mean accuracies and confidence interval for classification performed using 2-, 3-, and 5-classes models.

In order to estimate the expected classifiers' performance, we used a generalized linear model with a binomial link, as shown in the equation below.

$$\log\left(\frac{P}{1-P}\right) = \alpha + \beta_1 C_1 + \cdots + \beta_{n-1} C_{n-1}$$

where P is the accuracy of classification in each of the cross-validation runs, n is the number of groups/classes, α is the intercept representing the log odds-ratio of correct classification of the easiest to classify group, and β_1 to β_{n-1} are the coefficients of the model showing the increase (or decrease) in the classification odds. C_1 to C_{n-1} encode the remaining classes. The table shows estimated marginal means, standard errors, and lower and upper confidence limits.

Classifier	Species	Accuracy	SE	Lower CL	Upper CL
2-class	*Gyromitra*	0.986	0.00272	0.979	0.990
2-class	*Morchella*	0.983	0.00295	0.976	0.988
3-class	*Clavulina*	0.987	0.00188	0.982	0.990
3-class	*Inocybe*	0.967	0.00296	0.960	0.972
3-class	*Marasmius*	0.970	0.00281	0.964	0.975
5-class	*Agaricus*	0.973	0.00346	0.966	0.979
5-class	*Amanita*	0.982	0.00248	0.976	0.986
5-class	*Cantharellus*	0.984	0.00194	0.980	0.988
5-class	*Pleurotus*	0.979	0.00242	0.974	0.983
5-class	*Tricholoma*	0.958	0.00314	0.951	0.963

Figure S5. Comparison of the class accuracies in three different classifiers computed using 10× cross-validation. The red box indicated the mean, and lower and upper confidence intervals. Each point represents a single run of randomized cross-validation.

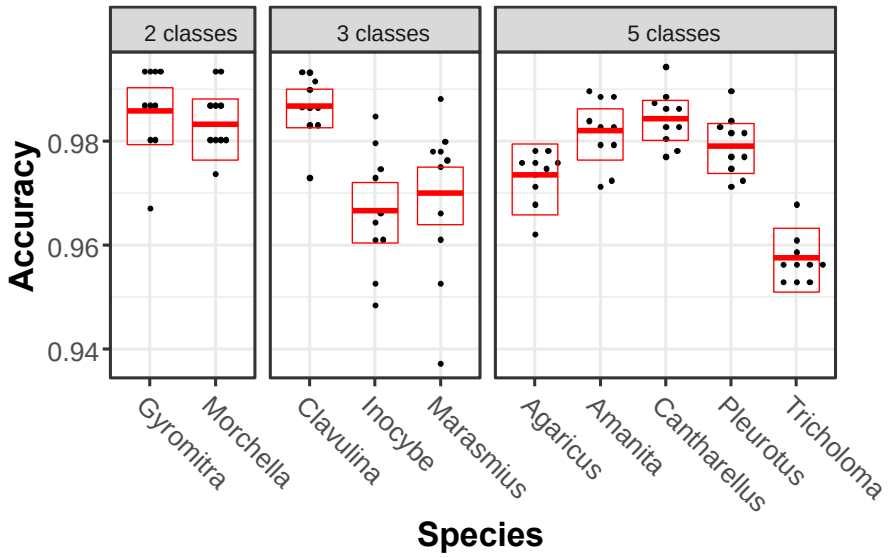


Table S2. Significance test results.

Classifier	Comparison	Odds ratio	SE	p-value
2-class	*Gyromitra* vs. *Morchella*	1.188	0.313	0.512
3-class	*Clavulina* vs. *Inocybe*	2.562	0.437	<0.001
3-class	*Clavulina* vs. *Marasmius*	2.305	0.399	<0.001
3-class	*Inocybe* vs. *Marasmius*	0.899	0.120	0.705
5-class	*Agaricus* vs. *Amanita*	0.675	0.131	0.254
5-class	*Agaricus* vs. *Cantharellus*	0.584	0.107	0.029
5-class	*Agaricus* vs. *Pleurotus*	0.784	0.140	0.655
5-class	*Agaricus* vs. *Tricholoma*	1.629	0.252	0.014
5-class	*Amanita* vs. *Cantharellus*	0.865	0.163	0.938
5-class	*Amanita* vs. *Pleurotus*	1.162	0.213	0.925
5-class	*Amanita* vs. *Tricholoma*	2.412	0.386	<0.001
5-class	*Cantharellus* vs. *Pleurotus*	1.344	0.232	0.427
5-class	*Cantharellus* vs. *Tricholoma*	2.791	0.412	<0.001
5-class	*Pleurotus* vs. *Tricholoma*	2.076	0.293	<0.001

The above-described model was used to compute comparisons between the classification accuracies achieved by the three evaluated classifiers. For multiple comparisons, the post-hoc statistical analysis utilized Tukey adjustment. The magnitude of the odds ratio reflects the strength of the relationship between the s and the proportion of correctly classified images. In the 2-class classifier, the results indicate that the expected accuracies for both classes are likely

to be identical. *Clavulina* is significantly easier to classify than the other two classes using the 3-class classifier. The 5-class classifier produces significantly less accurate results for the *Tricholoma* class, and significantly more accurate results for the *Cantharellus* class. Even though there are statistically significant differences in performance shown for some classes, the fact that even the worst noted accuracy exceeded 95% demonstrates that these differences may not have practical implications.