

COMPARATIVE ANALYSIS OF MODERN ONLINE TOOLS FOR OBJECT RECOGNITION IN IMAGES

Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine

There are many practical tasks that require object recognition in images. Various approaches are used to solve this task, typically involving neural networks [1]. These networks have been pre-trained on a specific set of reference images, with corresponding objects assigned to specific classes.

In this work, an analysis of 4 online tools providing functionality for object recognition in images has been conducted. The tools are as follows: Amazon Rekognition, Google Cloud Vision, Microsoft Azure AI Vision Studio, and Imagga. To investigate and assess their performance, the graphical interface of each tool's website was used. It is worth adding that these services allow users to interact with them both through their graphical interfaces and via API.

To analyze the quality of object recognition, 8 images were selected. These images cover a wide range of classes and exhibit diversity. The types of images can be described by the following list: street photo, interior, exterior, still life, animals, and sports (Figure 1).



Figure 1. The testing images

The recognition results of the «still life» image type for each tool are shown in Figure 2.

To calculate the metric for evaluating recognition quality, we will apply the following approach [2]. For each tool and each image, we will limit the size of identified objects to the top 20 in descending order of the system’s confidence in the accuracy of the prediction. Confidence is represented as a percentage value from 0 to 100. We will analyze the recognition quality as follows: for a correct prediction we add the confidence value. And for an incorrect prediction we subtract it. The resulting sum is divided by the number of recognized objects. The final assessment will be the sum of the tool’s results across all images divided by the number of images:

$$MeanQuality = \frac{1}{8} \sum_{i=1}^8 \frac{1}{L_i} \sum_{j=1}^{L_i} P_{ij} \cdot sign(P_{ij}), \quad (1)$$

where L_i denotes the number of recognized objects for the i -th image;
 P_{ij} denotes the confidence value of the tool for the j -th object of the i -th image;
 $sign(P_{ij})$ is a function that takes a value of 1 in the case of correct recognition and -1 in the case of wrong recognition.

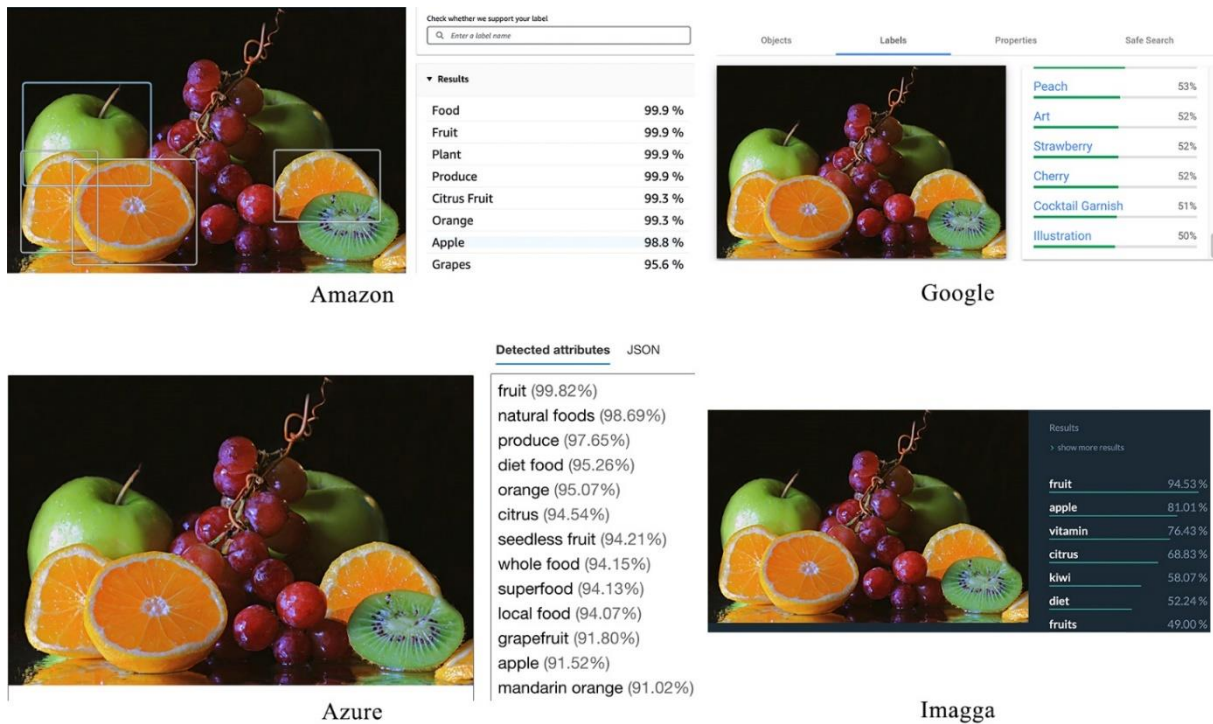


Figure 2. Recognition results by online tools

The calculated metric is presented in Table 1. The higher the metric value, the higher the tool’s rating.

Amazon and Azure are considered the best. Still life turned out to be the most challenging for all, except Imagga (we don’t see any meaningful drop). However, Imagga makes significant errors for images like Street photo 1 and Animals, which are easy for other tools. This may indicate that Imagga has been trained using different

methods and/or datasets. Such characteristics of Imagga show that it could be considered as a candidate for inclusion in ensemble models. However, this hypothesis needs further verification.

Table 1 – Results of the recognition quality analysis

Image	Amazon	Google	Azure	Imagga
Street photo 1	88,46	43,65	67,78	8,40
Street photo 2	33,43	72,35	63,02	19,35
Interior 1	55,46	9,8	66,45	28,04
Interior 2	67,29	8,7	85,56	31,83
Exterior	54,27	64,25	49,33	23,54
Sports	34,61	52,05	76,46	22,94
Animals	55,28	44,9	70,03	12,70
Still life	18,86	-5,6	27,80	26,54
Mean Quality	50,96	36,26	63,30	21,67

In addition, let us visualize the correctness of the detection for each tool based on all images. By x-axis we show the sum of confidence value for correct detections, by y-axis – the sum of confidence value for false detections. Each point on the chart represents the result for one image (Figure 3).

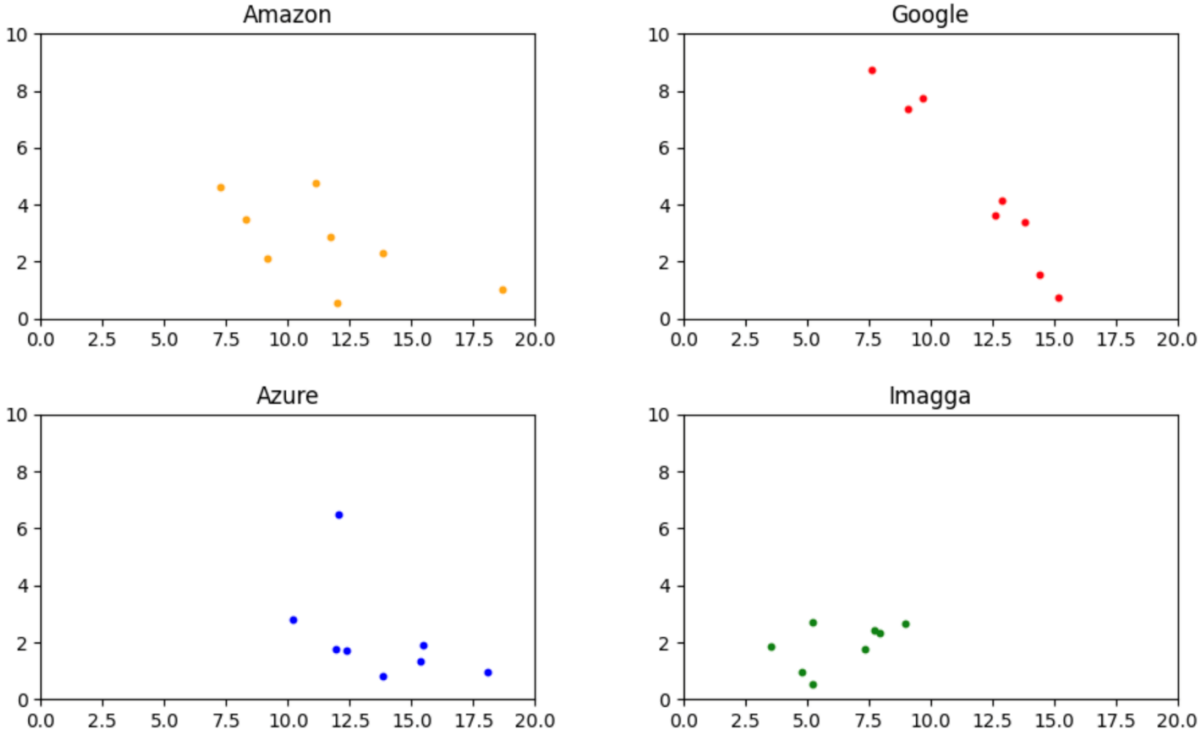


Figure 3. Visualization of the detection correctness for all tools

As we can see, all tools except Google have pretty tight scattering. This indicates that the quality of the detection does not vary significantly from one image to another.

In contrast Google scattering is linear which means that the quality of the detection depends on the image and is not as stable as observed with other tools. Azure scattering is located in the lower right corner which represents its highest ranking in current analysis.

References

1. Krizhevsky, A., Sutskever, I., Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90.
2. Rachev, S. T., Stoyanov, S., Fabozzi, F. J. (2008). Advanced stochastic models, risk assessment, and portfolio optimization: The ideal risk, uncertainty, and performance measures. *Optimization* (p. 382).

УДК 681.3.06

*Бобошко В. В., здобувач 2 курсу спеціальності 122 Комп'ютерні науки,
Штовба С. Д., д-р техн. наук, професор, професор кафедри інформаційних
технологій*

МЕТРИКА СХОЖОСТІ НАВЧАЛЬНИХ ДИСЦИПЛІН ЗА ЇХ ВНЕСКОМ У КОМПЕТЕНТНОСТІ

Донецький національний університет імені Василя Стуса, м. Вінниця

Після прийняття в 2015 р. нового закону «Про вищу освіту» освітня діяльність в університетах та в інших закладах вищої освіти розпочала формуватися навколо освітніх програм. Освітня програма – це документ, що визначає змістовний та процедурний складники навчання студентів та інших здобувачів вищої освіти. Освітні програми розроблюють за компетентісним підходом з урахуванням вимог освітніх стандартів за спеціальністю. Освітні стандарти є доволі загальними – вони містять перелік обов'язкових компетентностей та програмних результатів навчання. Натомість освітні програми, окрім компетентностей та програмних результатів навчання, містять і перелік освітніх дисциплін, обсяг кредитів та матриці впливу освітніх компонентів на програмні результати навчання й на компетентності.

Метою дослідження є розробка метрики для оцінювання схожості двох дисциплін з освітніх програм з однієї і тієї ж спеціальності. Метрика принципово відрізняється від традиційного суб'єктивного порівняння змісту дисциплін освітніми експертами. Метрика не враховує змістовне наповнення дисциплін, а розраховує схожість на основі внеску дисципліни у формування компетент-