2022

Seek, Picture Insect, Google Lens: An Analysis of Popular Insect Identification Apps Using Photos of Realistic Quality

Morgan Manderfield

Seek, Picture Insect, Google Lens: An Analysis of Popular Insect Identification Apps Using Photos of Realistic Quality

Morgan Manderfield
In partial fulfillment of the requirements for Master of Science, Entomology
University of Nebraska-Lincoln
Fall 2022

Abstract

Photo-based identification apps for mobile devices have grown in popularity across many fields of organismal biology. While studies have shown that these apps show promise in fields such as botany and ornithology, there is much skepticism surrounding their accuracy and usefulness in entomology. While providing these apps with photos of high quality will undoubtedly lead to the highest accuracy rates, often users submit photos of average quality, e.g., in poor lighting, out of focus, from far away, etc. To quantify realistic accuracy rates of photo-based insect identification apps, this study analyzed three popular apps – Seek, Picture Insect, and Google Lens – using photos of variable quality to simulate realistic/typical use. Taxonomic specificity of app results was also quantified and used to determine relative usefulness of each app. Seek produced the highest accuracy rates of the three, but often yielded taxonomically obvious results, resulting in low usefulness rates across all photo quality tiers. Picture Insect and Google Lens produced moderately high accuracy and usefulness rates for photos of ideal quality, but ≤54% accuracy for photos of acceptable and poor quality.

Introduction

There is no shortage of mobile phone apps available for download which center on exploring the natural world. In a 2015 search of Google Play Store apps, 6,301 nature-themed results were returned (Jepson and Ladle 2015), and since then many more have been developed and released. Growing in popularity among these are photo-based identification apps, designed to help users identify plants, mushrooms, birds, insects, and even rocks. Identification apps have many possible uses and reach diverse audiences – from teachers and students employing a study tool, to hobbyists attempting to self-teach, to professionals, citizen scientists, and academics seeking specimen identification for their work or research.

One emerging concern regarding arthropod identification apps in particular is the observation that they sometimes inaccurately suggest medically significant and/or emotionally charged identifications for innocuous species. Because insect identification apps are commonly used by people with little or no identification experience/insect knowledge, this can have the consequence of sending users into a panic over a harmless arthropod that, according to the app, is a parasite, a highly pestiferous or invasive species, or significantly venomous. People in general notoriously struggle to accurately identify arthropods, including medically significant species. In a survey of 2000 Americans commissioned by a developer of consumer insecticide products, it was revealed that although 62% respondents claimed to be confident in their ability to identify common insects, only 40% were able to correctly identify a bed bug, and 24% couldn't distinguish a wasp from a honey bee (Zevo 2019). Another study found that only 13% of 391 surveyed Germans were able to recognize a bed bug (Seidal and Reinhardt 2013). Yet another survey found that among 640 questionnaire respondents across the United States, the accuracy rate for correctly identifying common stinging Hymenoptera was 53% (Baker et al. 2014). It may seem to be a perfect solution then to have apps be able to do the heavy lifting in identifying insects for laymen. However, given that most people do not possess sufficient knowledge to be able to distinguish accurate from inaccurate results, the accuracy rates of these apps would need to be extremely high to minimize the likelihood and prevalence of incorrect responses being taken at face value.

Reports on accuracy of mobile photo-based identification apps are variable. A Rutgers University analysis of 6 plant identification apps found that PictureThis and iNaturalist could identify tree species reasonably well at the genus level but struggled at the species level (PlantSnap failed to consistently produce accurate results). The authors suggest that the most appropriate use of these apps is to guide unsure arborists toward a correct identification by refining the pool of possible species to be further investigated (Schmidt et al. 2022). When assessing apps for their ability to identify toxic plant species, researchers from University of California San Diego and University of Colorado Boulder determined a 59% success rate for PictureThis but a dismal 5.8% success rate for PlantSnap (Otter et al. 2020). Researchers at the Junior Academy of Science of Ukraine sought to quantify the ability of Google Lens to identify plants when provided with photos of variable quality; they found that even at relatively low resolution/poor quality, the app performed surprisingly well with a failure rate of 14.3% (Shapovalov et al. 2020).

The inconsistency in the ability of mobile apps to accurately identify various flora and fauna has made some users and would-be users skeptical. A poll conducted by the author within the Sci-comm Facebook group “Entomology” (which serves a diverse, global membership base ranging from insect novices to world renowned professional and academic entomologists) revealed polarized opinions among insect enthusiasts and scientists surrounding use of photo-based insect identification apps. Of the 178 respondents, exactly 50% claimed that they do not use arthropod ID apps, while the other 50% claimed that they do, with several respondents claiming to use multiple apps for arthropod identification (Manderfield M., unpublished data 2022).

Another source of skepticism surrounding the ability of apps to accurately identify insects is the sheer number of insect species that exist. Compared to approximately 450,000 plant species worldwide (Pimm and Joppa 2015), and approximately 18,000 bird species worldwide (Barrowclough et al. 2016), there are an estimated 900,000 described insect species and potentially tens of millions of unique insect species on the earth (Smithsonian Institution 1996).

Since 2018, three photo-based insect identification apps have been released which seem to overshadow all others in popularity and positive user reviews: Seek (iNaturalist), Picture Insect (Next Vision Limited), and Lens (Google). Seek claims to be able to identify insects as well as other wildlife, plants, and fungi, drawing from millions of observations posted to iNaturalist by users. Picture Insect was developed for the purpose of insect and spider identification and claims to be able to identify more than 4,000 species with an accuracy rate of >95%. While Lens is not intended to strictly be an insect identification app, it is widely used for that purpose, and identification of plants and animals is listed among the developer claims. Likely due to their recent rise to popularity, few studies have been conducted to analyze the effective accuracy rates of these apps in identifying insect and arachnid species. Furthermore, few studies have been conducted which analyze the efficacy of photo-based identification apps in general for their ability to perform in less-than-ideal conditions. How accurate are these apps when provided with photos of out-of-focus subjects? Photos with subjects that are perhaps too far away, at an awkward angle, or in poor lighting? Simply put, how do these apps perform when provided with realistic photos, taken in the field by amateurs who know little about their subjects? The investigation presented herein seeks to answer these questions by analyzing and comparing the accuracy rates of these three apps when provided with photos of variable quality. Furthermore, these accuracy rates are put into perspective by quantifying taxonomic specificity of identifications and reported on as rates of usefulness of results.

Materials & Methods

App selection

Three apps were evaluated: Seek (iNaturalist), Picture Insect (Next Vision Limited), and Lens (Google). Each app was installed to a Pixel 5a device for use during the evaluation. The most recent version of each app at the time was used, and the apps were confirmed to be recently updated prior to being used (Table 1).

	No. of Downloads	Version	User Review Rating	Release Date	Last Updated on Pixel 5a
Picture Insect	>1M	2.8.6	4.3/5	Oct 18, 2019	Sept 19, 2022
Seek	>1M	2.14.2	4.5/5	Jan 18, 2019	July 22, 2022
Lens	>1B	1.14.220323019	4.6/5	Jun 4, 2018	May 11, 2022

Table 1: App information at the time of evaluation

Photo sourcing and categorization

Each photo collected was placed into one of three quality categories/tiers: Ideal, Acceptable, or Poor. Photos of Ideal quality met all the criteria generally recommended by the apps: they were zoomed in on the specimen, the specimen was in focus, lighting was good, the specimen was viewed from a good angle, and there was minimal background noise (Figure 1a). Acceptable photos generally met these criteria but may have been lacking in one tier, such as being on a somewhat noisy background or viewed from an imperfect angle (Figure 1b). Poor photos were either strongly lacking in one criterion or lacking in several criteria. However, they were deemed insightful enough that an app may have a reasonable chance at producing a useful identification (Figure 1c). Photos of Acceptable and Poor quality were included in an attempt to realistically simulate the imperfect photos that are frequently submitted by app users.

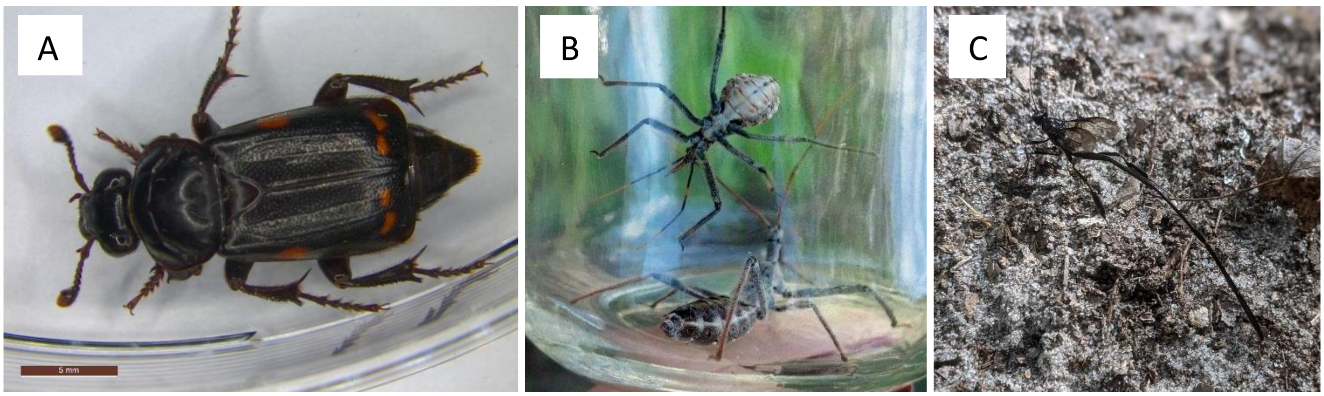

Figure 1: Examples of photos belonging to each quality category. A = Ideal quality (*Nicrophorus pustulatus*), B = Acceptable quality (*Arilus cristatus*), C= Poor quality (*Pelecinus polyturator*)

A total of 100 photos were used, sourced from multiple entities. 35 of the photos were from the author's personal collection (n=29) or provided by acquaintances (n=6). The remaining 65 were sourced from a commercial pest control company where the author is employed; these photos were submitted by service technicians to the company's research and development team for specimen identification and subsequently saved in a repository. These 65 photos were mostly of Acceptable and Poor quality and best represent the type of photos that would be commonly run through identification apps (taken quickly, in the field, perhaps in non-ideal conditions, by amateurs/non-entomologists, etc.). Of the 100 photos, 32 were categorized as Ideal, 37 were Acceptable, and 31 were Poor. A total of 16 insect or arachnid orders were represented among the photos. 84 of the photos depicted adults, while 16 of the photos depicted immature forms –

either larvae, nymphs, or pupae. Most of the photos were taken in the continental United States, except 3 which were taken in Canada and 6 in Puerto Rico (depicting species with a cosmopolitan distribution).

Subject identity establishment

Prior to being run through the apps, subject identity was determined for each photo. Identity was established in one of two ways. Photos owned by the author were hosted on the iNaturalist website and were used as part of the study only if they achieved "Research Grade" status (location and date information are provided, and identity is corroborated by a minimum of 2 iNaturalist contributors, many of whom are professional entomologists or insect specialists). For photos that could not be hosted on iNaturalist due to confidentiality concerns, subjects were still identified by a minimum of 2 professional entomologists with extensive experience in insect identification, including the author.

All photo subjects were identified to at least the family level. 70 subjects were identified to species, 12 to genus, 4 to subfamily, and 14 to family.

Determination of accuracy

Each photo was run through each app one time and the result was recorded as a unique event. An app result was considered accurate if it produced a taxonomically correct identification, regardless of classification level (e.g., if the result "Animalia" was given for a photo of a cockroach, this was recorded as accurate). Frequently, the apps provided multiple results for a single photo; if the primary result was inaccurate but the secondary options contained an accurate result, the event was recorded as accurate. If the result given by the app was inaccurate, it was recorded as such regardless of how "close" it was to being accurate (e.g., if the result "*Amblyomma cajennense*" was given for a subject previously established to be *Amblyomma maculatum*, the event was recorded as inaccurate with no additional remarks).

Determination of usefulness

In an attempt to contextualize taxonomically obvious results (e.g., Animalia, Arthropoda) being recorded as accurate and potentially boosting an app's overall perceived accuracy, each accurate event was also determined to be useful or not. An event was determined to be useful if it produced an accurate identification at the family level or lower. This definition was chosen under the assumption that most people use insect identification apps intending to learn about the photo subject's biology and behavior, and generally, biology and behavior varies too significantly in higher taxonomic levels to be of usefulness to the app user.

If an app produced a primary result that was not useful (higher than family level), but a secondary result that was useful (family level or lower), the event was recorded as useful. Usefulness was not recorded for events which produced exclusively inaccurate results.

Determination of overconfidence

Some photo subjects were established to only be identifiable to a certain taxonomic level above species (e.g., family, subfamily, or genus), due to either photo quality, developmental/life stage, or necessity of examining the specimen in person, under a microscope, from a different angle, etc. If an app produced an identification that was more specific than what was previously established to be reasonably possible, the event was recorded as being overconfident in lieu of accurate or inaccurate.

Data analysis

Accurate, useful, and overconfident results were totaled for each app and averages were calculated to determine accuracy, usefulness, and overconfidence rates. Accuracy and usefulness rates were also determined within each photo quality tier for each app. Usefulness rate was determined per total accurate results, rather than per total photos.

Results

Lens was unable to render a result for 9 photos of 100 (9%). In these situations, instead of offering an identification, Lens presented "visually similar images" from Google Image search results to the user. Of the 9 photos that failed to produce a result, 2 were Poor, 5 were Acceptable, and 2 were Ideal. Picture Insect and Seek produced a result 100% of the time.

Overall accuracy rates for Seek, Picture Insect, and Lens can be seen in Figure 2. Seek achieved the highest overall accuracy rate (95%), followed by Lens (57%) and Picture Insect (53%). Included in these rates are results which produced an inaccurate primary identification but an accurate secondary identification.

Accuracy rates for each app broken down by photo quality tier can be seen in Figure 3. When looking exclusively at results for Ideal photos, Seek achieved 97% accuracy, and accuracy rates for Picture Insect and Lens raised significantly from their overall average to 72% and 78% respectively. Seek achieved consistently high accuracy across all photo quality tiers, with the lowest accuracy rate being 90% for Poor photos. For each app, accuracy rates increased as photo quality increased (except for Seek, which saw the same accuracy rate for Acceptable and Ideal photos).

Overconfidence rates for Seek, Picture Insect, and Lens were 1%, 17%, and 3% respectively. Overall usefulness rates were 41%, 98%, and 91% respectively (Figure 2).

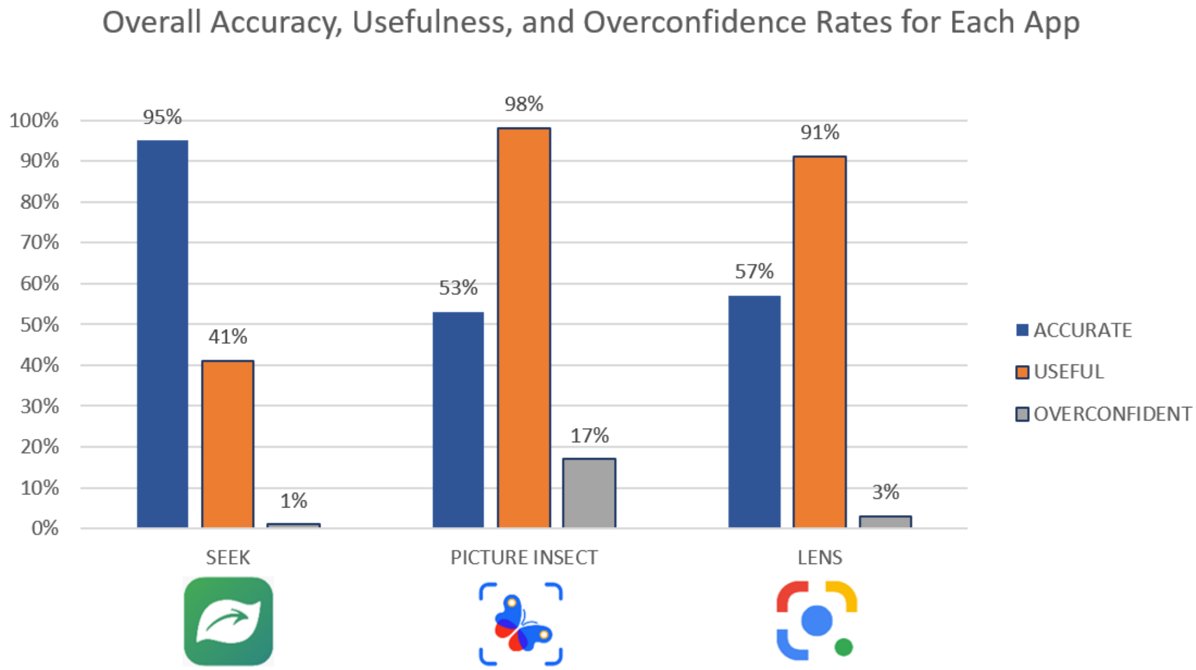

Figure 2: Chart depicting accuracy, usefulness, and overconfidence rates for each app. Usefulness rate is per total accurate results, rather than total photos. Rates calculated from combined results from all photo quality tiers.

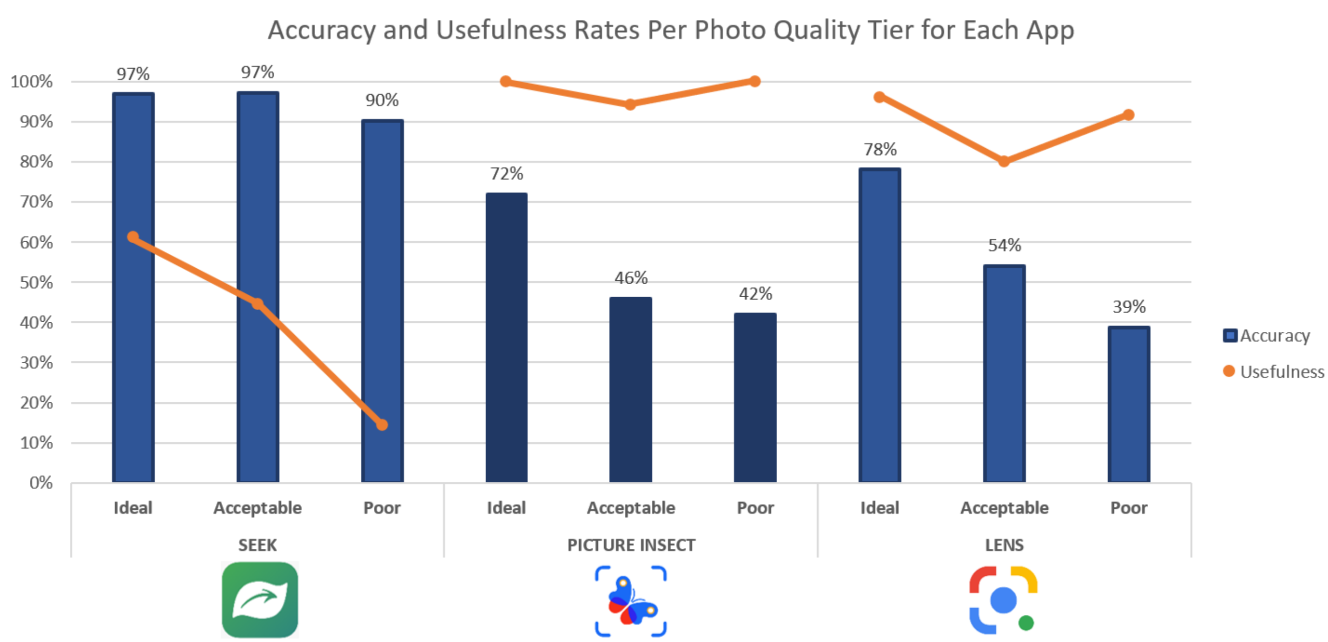

Figure 3: Chart depicting accuracy rates for each photo quality tier for each app (bars), and usefulness rate per quality tier for each app (lines). Usefulness rate is per total accurate results, rather than total photos.

When broken down by photo quality tier, the usefulness rates for each app were as follows. For Seek: 61% (Ideal); 44% (Acceptable); 14% (Poor). Picture Insect: 100% (Ideal); 94% (Acceptable); 100% (Poor). Lens: 96% (Ideal); 80% (Acceptable); 92% (Poor) (Figure 3).

Discussion

Accuracy and Usefulness

The positive correlation observed between photo quality and accuracy rate is to be expected; the apps themselves instruct users to take clear, magnified photos in good lighting to increase the likelihood of receiving a more accurate result.

Despite the initially impressive accuracy rate for Seek across all photo quality tiers, it is worth noting that many of these identifications were taxonomically obvious. 34 of 100 results for Seek were at the subclass level (e.g. Pterygota, all winged and secondarily wingless insects) or higher, and 13 results were at the subphylum level (e.g. Hexapoda, which includes all insects as well as Collembola, Diplura, and Protura) or higher. This contributed significantly to the relatively low usefulness rate for Seek; Seek's highest usefulness rate, for Ideal photos, was still well below the usefulness rate for even Acceptable and Poor photos run through the other apps.

While Seek tended to skew toward higher classification results, Picture Insect produced more specific identifications. A species-level identification was given by Picture Insect 99 out of 100 times. This had a significant impact on the overall usefulness rate of Picture Insect; nearly every time the app produced an accurate result, it was also useful. The one occasion that Picture Insect produced an identification higher than species level was for a photo of Phoridae (Diptera) pupae, which the app recognized as belonging to Cyclorrhapha, un unranked taxon within the Muscomorpha infraorder.

Lens and Picture Insect followed similar accuracy/usefulness trends as one another. While accuracy rates for Lens were slightly higher than Picture Insect for Ideal and Acceptable photos, it was slightly lower for Poor photos. Lens also achieved relatively high usefulness rates, but less than Picture Insect when comparing each photo quality tier individually. Interestingly, both Picture Insect and Lens achieved the lowest usefulness rating for Acceptable photos, whereas the lowest usefulness rating for Seek was for Poor photos (by a rather large margin).

One noteworthy observation was the prevalence of "favorite" or recurring identifications for a few of the apps, as well as recurring failure to identify certain taxa. For example, only 1 of the 4 times Picture Insect offered *Oryzaephilus surinamensis* was it accurate (the three other submissions were a different beetle from Silvanidae and two beetles from Latridiidae). Picture Insect similarly struggled to identify Psocodea; 2 out of 3 times when Ideal and Acceptable photos of Psocodea were submitted, the app offered a termite species instead (the 3rd time, a Collembolan species was offered). Picture Insect also struggled with Phoridae (Diptera), offering instead *Drosophila melanogaster* 3 out of the 5 times this taxon was submitted.

On several occasions, the apps inaccurately identified innocuous species as a medically significant species. This was observed once for Picture Insect and several times for Lens. Picture Insect identified a female *Thylodrias contractus* (Dermestidae) as a bed bug, *Cimex lectularius* (Cimicidae). Lens offered the identification of "bed bug" incorrectly as a primary result twice, once for *Adistemia watsoni*, once for *Eufallia sp.*, both members of Latridiidae. Lens also offered "bed bug" in secondary results numerous times, however these were not quantified. Lens identified another Latridiid beetle, *Cartodere constricta*, as a parasitic louse. On another occasion, Lens correctly offered a primary result for an Acceptable photo of an Eastern cicada killer, *Sphecius speciosus*, but as a secondary result offered *Vespa mandarinia* (Vespidae). This is especially noteworthy because of the country-wide, media- and social media-fueled panic surrounding *V. mandarinia* that unfolded in the United States during the spring and summer of 2020, when an introduction of the species to Washington state resulted in a massive uptick in wasp identification requests sent to university entomology departments (Skvarla et al. 2022) and sweeping misidentification and killing of native wasps and wasp lookalikes by the public (Marantos 2020, Russell 2022). Users of insect identification apps who lack sufficient entomological knowledge to be able to differentiate these species themselves are likely to internalize startling/upsetting app results, even when offered as a secondary result.

Overconfidence

Overconfidence rates for Seek, Picture Insect, and Lens were 1%, 17%, and 3% respectively. The very low overconfidence rate for Seek makes intuitive sense, because Seek tended more toward less specific/higher classification identifications. Picture Insect had a relatively high overconfidence rate by comparison, which can easily be explained by the app's tendency to provide species-level identifications. A species identification was pre-determined for 70% of photo subjects, while the other 30% were left at genus, subfamily, or family. It is mathematically fitting that overconfidence rate would be high, considering that a species-level identification was given 99% of the time.

App user experience

The overall app user experience and unique "quirks" of each app are worth discussing here, some of which may have an impact on accuracy and usefulness yet are somewhat difficult to quantify. For example, Picture Insect and Lens allow the user to crop the photo/zoom in on the photo subject within the app itself but Seek does not. An outcome of this is that many times, a more or less accurate result can be rendered for Picture Insect and Lens simply by adjusting the area of the photo for the app to focus on. By contrast, Seek conservatively opts for an accurate but taxonomically obvious result and instructs the user to crop the image and try again, which is arguably cumbersome and something that users may ignore or be dissuaded by.

Because Seek serves as an identification app for many types of fauna and flora, it offers a filter option which users can activate for “Plants only” or “Non-plants only.” The “Non-plants only" filter was turned on for this evaluation. It’s possible that this option gives Seek an advantage over Lens, which simply compares photos to visually similar images in the Google Image repository, and does not distinguish between plants, animals, or other. Since Picture Insect is designed as an insect/arachnid identification app only, it seems unlikely to consider or offer non-insect/arachnid results, though this could not be confirmed. On one occasion, Lens did offer a non-arthropod identification as a result, identifying a *Hyalophora cecropia* (Saturniidae) larva as a pumpkin (Figure 4). This photo was included as a stretch goal for the apps, since the specimen itself is in poor/atypical condition (deceased, discolored, possibly parasitized). Both Seek and Picture Insect were able to accurately identify this subject, with results of “Lepidoptera” and “*Hyalophora cecropia*,” respectively.

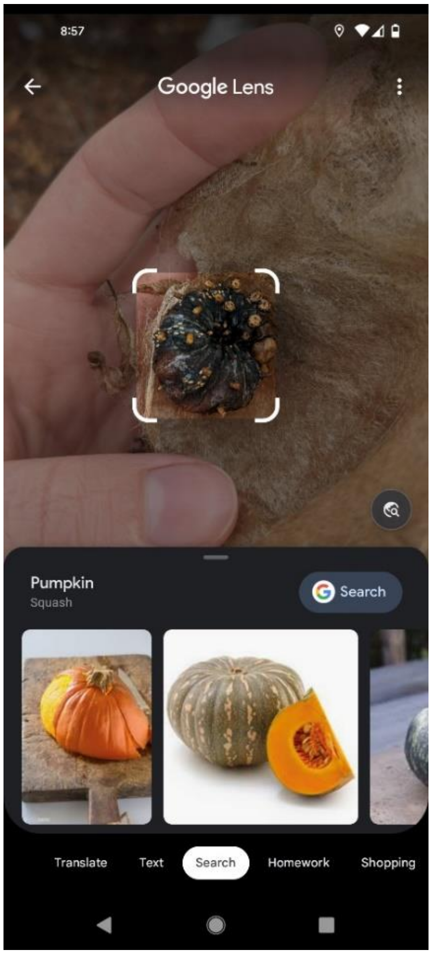

Figure 4: Lens result for a photo of a deceased, possibly parasitized *Hyalophora cecropia* (Saturniidae) larva, suggesting an identification of "Pumpkin"

Occasionally, the apps produced regionally inappropriate identifications (e.g. species occurring exclusively in Europe or Asia). This did not occur with Seek (likely because Seek skewed more toward higher taxonomic classifications rather than species). It occurred once for Picture Insect and occurred 9 times among primary identifications for Lens (rate of occurrence among secondary identifications was not recorded). Seek does offer an option for the user to input their locality into their app profile with the claim of being able to offer region-specific identification suggestions. This feature was not used for this evaluation, because photos from across many locations were evaluated. Interestingly, Picture Insect does not offer a feature like this, but it produced regionally appropriate results in all cases except one. One possible explanation for this is that of the 4,000 insects that Picture Insect claims to be able to identify, many or all of them occur in North America. The one regionally inappropriate result that Picture Insect provided was *Gibbium psylloides* (Ptinidae) for a photo of *Gibbium aequinoctiale*; this error is unsurprising and perhaps even forgivable considering *G. aequinoctiale* is somewhat frequently misreported as *G. psylloides* (ISU 2022). The higher occurrence of regionally inappropriate results produced by Lens is likely explained by how the app works, simply looking for similar images within the Google Images repository and not taking geography into account. Furthermore, if images are mislabeled/misclassified by the photo owner or host website, they may be erroneously included in the pool of photos which Lens assigns a certain identification to.

Limitations

Certainly, a larger data set would have been desirable to be able to evaluate how the apps process photos of subjects from more/different taxonomies, more life stages, from more geographies, etc. Acquiring high quality photos of insect specimens is not difficult, as there is an abundance of such photos to select from. However, as one of the objectives of the study was to simulate realistic use of the apps, photos of mediocre and poor quality were needed. Such photos are in general less readily available because photographers tend not to publish them to publicly available photo repositories (e.g. iNaturalist, bugguide, Flickr, etc.). Additionally, requiring photos of poor enough quality to be categorized as such but high enough quality that the subject identity can still be independently verified is further limiting.

It also must be acknowledged that the hand-picked nature of the photos used in this evaluation likely had an influence on the results. To be able to confidently verify app accuracy, photo subject identity needed to be independently verifiable by two entomologists. However, because resources to hire taxonomists/specialists were not available, subjects were voluntarily identified by iNaturalist contributors, professional entomologists within the author's social sphere, and the author. As a result, many of the photo subjects were distinctly patterned/colored or otherwise "obvious" in identity. In reality, a vast number of insects are relatively nondescript (e.g. Microlepidoptera), and many of them even require dissection to accurately identify. Photos of such species could not be included in the study because their identities were not verifiable. It's plausible that if professional entomologists could not identify these photo subjects (easily or at all), apps would struggle to accurately identify them as well.

Ideally, professional/expert taxonomists would have been utilized to verify photo subject identity. This would have resulted in greatly increased diversity of photo subjects that were selected to be used in this study. While great care was taken by the author and professional entomologist peers to ensure overconfident identifications were not given, the same level of certainty cannot be guaranteed from iNaturalist contributors. It is true that many iNaturalist contributors are professional taxonomists or otherwise greatly skilled at arthropod identification. However, the credentials and expertise of each iNaturalist contributor who provided subject identification for this evaluation could not be verified.

The method for determining overconfidence may also have impacted accuracy/usefulness rates. It is improbable, but certainly possible that in situations where a result was determined to be overconfident, the app in fact rendered an accurate result, yet the entomologist identifiers did not possess adequate knowledge to provide a more specific identification.

Nearly all photos used for this study were taken in the United States and Canada, and those taken elsewhere (Puerto Rico) depicted species with cosmopolitan distributions. Due to the limited geographic range the data set spans, conclusions from this study are only relevant to North American insect/arachnid species, specifically the United States and Canada. This is certainly limiting, as these apps are available and undoubtedly used in other geographies; it's unknown how these apps perform when subjects from outside of North America are submitted.

Suggestions for future study

There is ample opportunity to expand upon this investigation and delve deeper into this subject matter for a more comprehensive understanding of arthropod identification app efficacy, potential areas of app improvement, use patterns, impact, etc. To draw directly from the limitations of this investigation, expanding the geographic range from which photos are collected from would certainly allow for a better understanding of how these apps (and others) perform across the world. This is valuable information because many insect identification apps, including those evaluated herein, are available to global communities and some claim to draw upon data from all over the world.

The ability to accurately identify medically significant species and their frequently mis-identified "lookalikes" is an important success criterion for arthropod identification apps. Fear and distrust of arthropods remains extremely pervasive (Zevo 2019), and users turn to identification apps either for reassurance that their specimen is harmless or to arm themselves with accurate information for how to approach a potentially dangerous arthropod. This presents not only an opportunity, but arguably an obligation for app developers to hyperfocus on medically significant species to be sure they are delivering accurate results exactly when they are most needed and most sought after. The present study did include some medically significant species, but a more focused investigation into app efficacy specifically as it relates to positive identification of medically significant species (and successfully ruling out lookalike species) would be insightful and valuable.

Finally, detailed surveys to determine who the user base is, and the scope of app use would allow for a better understanding of impact. Who exactly are the users? Are they students, teachers, citizen scientists, insect enthusiasts, insect adversaries? Are they professionals, using the apps as part of their work? If so, how are the apps supporting that work? Knowing the answers to these questions not only helps app developers build a better product for their users, but also allows us to understand what the (cultural, environmental, academic, etc.) impact may be in situations where users take app identifications at face value.

Conclusions

Seek produced the highest accuracy rates (>90%) for all photo quality tiers, however due to frequency of taxonomically obvious results produced significantly lower usefulness rates for all photo quality tiers compared to the other two apps. Seek's "effective" accuracy rate (result is both accurate and useful) was 61% for Ideal photos, 44% for Acceptable Photos, and 14% for Poor photos. Picture Insect and Lens performed similarly to one another, achieving relatively high usefulness ratings for all photo quality tiers. Picture Insect and Lens achieved 72% and 78% accuracy for Ideal photos respectively but produced accuracy rates ≤54% for Acceptable and ≤42% for Poor photos.

The results of this study serve as a valuable benchmark for the effective/realistic accuracy rates of these popular apps. Accuracy rates were observed to decrease with photo quality for all three apps evaluated. While this is not groundbreaking or surprising, it's important to understand the limitations/precise failure points of these apps, specifically how they perform when supplied with photos of imperfect but realistic quality. Most photos used in this study were taken by amateurs/non-entomologists, in less-than-ideal conditions, with the intention to submit to an entomologist for identification. App developers are unlikely to incorporate these types of photos into their analyses, instead focusing on perfect use scenarios. Thus, higher accuracy rates are reported, and users may be given a false sense of confidence while using these apps. The attempt to quantify usefulness of app results is also valuable in allowing high accuracy rates inflated with taxonomically obvious/unhelpful results to be put into perspective.

Acknowledgements

This project received support from many entities, whom I would like to take the opportunity to acknowledge. I thank my colleague Alicia for identifying a huge number of photo subjects with the same degree of care and caution that I would have applied myself. I offer sincere thanks to all the individuals who enthusiastically contributed photos from their personal libraries to be included within the data set. I am also grateful for the volunteer contributors on iNaturalist who identified many of the photo subjects. I thank my friends and family for being my personal cheerleaders throughout this process. Finally, I acknowledge and thank my employer who provided access to so many photos of “mediocre” quality, which is truly a sincere sentiment because those were otherwise rather difficult to acquire.

References

Baker, T.W., Forester, J.P., Johnson, M.L., Stolfi, A., Stahl, M.C. 2014. The HIT study: Hymenoptera Identification Test—how accurate are people at identifying stinging insects? Ann. Allergy Asthma Immunol. 113(3): 267-270. DOI: https://doi.org/10.1016/j.anai.2014.05.029

Barrowclough, G.F., Cracraft, J., Klicka, J., Zink, R.M. 2016. How Many Kinds of Birds Are There and Why Does It Matter? PLoS ONE. 11(11): e0166307. DOI: https://doi.org/10.1371/journal.pone.0166307

(ISU). Iowa State University Department of Plant Pathology, Entomology, and Microbiology. 2022. Bugguide. Species *Gibbium aequinoctiale* – Smooth Spider Beetle. https://bugguide.net/node/view/55820

Jepson, P., and R.J. Ladle. 2015. Nature apps: Waiting for the revolution. Ambio. 44(8): 827-32. DOI: https://doi.org/10.1007/s13280-015-0712-2

Marantos, J. 2020. Panicked over 'murder hornets,' people are killing native bees we desperately need. Los Angeles Times. https://www.latimes.com/lifestyle/story/2020-05-08/panicked-over-murder-hornets-people-are-killing-the-native-bees-we-desperately-need

Otter J., Mayer S., and C.A. Tomaszewski. 2021. Swipe Right: a Comparison of Accuracy of Plant Identification Apps for Toxic Plants. J Med Toxicol. 17(1): 42-47. DOI: https://doi.org/10.1007/s13181-020-00803-6

Pimm, S. and L. Joppa. 2015. How Many Plant Species are There, Where are They, and at What Rate are They Going Extinct?. Annals of the Missouri Botanical Garden. 100: 170-176. DOI: https://doi.org/10.3417/2012018

Russell, A. 2022. It's big, but it's not a 'murder hornet'. Texas A&M AgriLife. https://agrilifetoday.tamu.edu/2022/07/28/its-big-but-its-not-a-murder-hornet/

Schmidt, R.J., Casario, B. M., Zipse, P. C., and J.C. Grabosky. 2022. An analysis of the accuracy of photo-based plant identification applications on 55 tree species. AUF. 48(1). DOI: https://doi.org/10.48044/jauf.2022.003

Seidel, C. and K. Reinhardt. 2013. Bugging Forecast: Unknown, Disliked, Occasionally Intimate. Bed Bugs in Germany Meet Unprepared People. PLoS ONE. 8(1): e51083. DOI: https://doi.org/10.1371/journal.pone.0051083

Shapovalov, V., Shapovalov, Y.B., and Z. Bilyk. 2020. The Google Lens analyzing quality: an analysis of the possibility to use in the educational process. AREdu. DOI:10.31812/EDUCDIM.V53I1.3844

Skvarla, M.J., Bertone, M.A., and P.J. Liesch. 2022. Murder Hornet Mayhem: The Impact of the 2020 Giant Hornet Panic and COVID-19 Pandemic on Arthropod Identification Laboratories. Am. Entomol. 68(2): 38-43. DOI: https://doi.org/10.1093/ae/tmac029

(Smithsonian Institution) Department of Systematic Biology, Entomology Section, National Museum of Natural History. 1996. Buginfo. Numbers of Insects (Species and Individuals). https://www.si.edu/spotlight/buginfo/bugnos#:~:text=In%20the%20United%20States%2C%20the,is%20estimated%20at%20some%2073%2C000

Zevo via OnePoll. 2019. Survey conducted by OnePoll for Zevo with a sample of 2,000 Americans from August 12-16, 2019. https://swns-research.medium.com/study-finds-that-men-are-more-likely-than-women-to-be-very-scared-of-creepy-crawlers-a2ad36a54217