Automatic Detection and Classification of Tick-borne Skin Lesions using Deep Learning

Lauren Michelle Pfeifer  
Independent Researcher  
San Francisco, CA 94107  
pfeifer.lauren@gmail.com

Matias Valdenegro-Toro  
German Research Center for Artificial Intelligence  
28359 Bremen, Germany  
matias.valdenegro@dfki.de

Abstract

Around the globe, ticks are the culprit of transmitting a variety of bacterial, viral and parasitic diseases. The incidence of tick-borne diseases has drastically increased within the last decade, with annual cases of Lyme disease soaring to an estimated 300,000 in the United States alone. As a result, more efforts in improving lesion identification approaches and diagnostics for tick-borne illnesses is critical. The objective for this study is to build upon the approach used by Burlina et al. by using a variety of convolutional neural network models to detect tick-borne skin lesions. We expanded the data inputs by acquiring images from Google in seven different languages to test if this would diversify training data and improve the accuracy of skin lesion detection. The final dataset included nearly 6,080 images and was trained on a combination of architectures (ResNet 34, ResNet 50, VGG 19, and Dense Net 121). We obtained an accuracy of 80.72% with our model trained on the DenseNet 121 architecture.

1 Introduction

Around the globe, ticks are the culprit of transmitting a variety of bacterial, viral and parasitic diseases (Rodriguez-Morales, et al., 2018). The incidence of tick-borne diseases, like Lyme disease, Tick-borne Encephalitis, and Rickettsiosis has drastically increased within the last decade \cite{10} \cite{3}. The number of annual cases of Lyme disease alone has soared to an estimated 300,000 in the United States \cite{6}. As a result, more efforts in improving lesion identification approaches and diagnostics for tick-borne illnesses is critical in preventing serious long-term complications \cite{7}.

The initial phase of our research specifically focuses on tick-borne skin lesions relating to Lyme disease, namely erythema migrans (EM), an erythematous, annular skin lesion \cite{1}. Erythema migrans emerges at the site of the tick bite as an expanding skin redness and often heals without antibiotic treatment, however the infecting pathogens can spread to tissues and organs throughout the body \cite{9}. Timely diagnosis and treatment of a tick-borne skin lesion, such as EM, can reduce the severity of tick-borne illness symptoms early on \cite{2}. For Lyme disease specifically, deep learning detection models have been applied to skin lesion (erythema migrans) images for the purpose of early diagnostics and treatment.

2 Methods and Experimental Results

Our study is conducting a binary classification of erythema migrans and non-erythema migrans. For non-erythema migrans images, we focused on images of mosquito bites and healthy skin. Mosquito bite images were selected due to the similar appearance in colour and texture (although different in size) from the typical “bulls-eye” shape of erythema migrans. People often mistake erythema migrans
as infected mosquito bites [3] which causes them to forgo getting proper treatment [2]. Similar to the
study of Burlina et al., binary classification is our end goal, however this paper focuses on a 3-class
classification approach of detecting erythema migrans, mosquito bite(s), and healthy skin.

As of now, there are still no publicly available datasets of tick bite images. For this study, the dataset
was compiled using images from Google. This approach was leveraged from fastai’s Jeremy Howard
and Adrian Rosebrock who use Google Images to gather training data [4] [8]. Our objective is to train
a classifier for three classes: tick bites, mosquito bites, and healthy skin, the latter which considers
skin without any kind of tick or mosquito bite.

Our initial search was originally conducted only in English using the search queries of “tick bite”,
“mosquito bite”, and “healthy skin” in Google Images. The results did not populate a diverse outcome
of images in terms of skin color, textures, and skin features. For this reason we decided to expand our
image data search to include six other languages: Spanish, French, German, Portuguese, Mandarin,
and Russian. The selection process for these languages included verifying geographies abroad that
have prolific cases of tick-borne diseases, which includes North America, Europe, Northern Asia, and
Latin America. We verified that the translations of each search query was correct by using Google
Translate. Our final dataset included 6,080 images.

We decided to keep our search queries less complex as a starting point to gather as much information
on tick bites as possible. In contrast, Burlina et al. used a combination of search terms (e.g. “Erythema
migrans”, “Lyme”, or “bullseye rash”) that are more relevant to Lyme disease [2]. In the next phase
of research we plan to fine tune our search queries based on the results we saw from our current
search results. Table 2 shows an overview of the queries we used for each language.

The key performance metric in the study was accuracy in demonstrating the number of predictions our
model correctly identified as a tick bite, mosquito bite or healthy skin. Out of the four architectures
we ran, our model obtained the best accuracy of 80.72% trained on the DenseNet 121 architecture as
indicated in Table 1. The second best result we obtained was an accuracy of 79.16% trained on the
ResNet 34 architecture. The remaining two architectures, resulted in lower accuracies with VGG-19
at 78.64% and ResNet 50 at 53.12%.

In addition to correctly classifying skin lesions, our results of our dataset showed to indeed be less
biased in terms of skin color, textures, and skin features.

Our study of using deep learning to automate detection and classification of tick-borne skin lesions
generated promising results. The incorporation of using diverse approaches to gather data, such as
our use of queries in multiple languages, exhibited a step in the right direction towards reducing
bias in skin colors. For instance, the search queries in Portuguese and Spanish for “tick bite” and
“mosquito bite” resulted in images on darker skin tones, while the search query for “healthy skin”
included more image results with darker skin tones as well.

As expected, we encountered some limitations from our broad scope of search queries. The results
for healthy skin exhibited more challenges in classifying healthy skin. For our next phase of research,
we aim to refine the search queries for healthy skin. For tick-borne skin lesions, we also plan to
expand the search queries and languages used to cover additional geographies and skin lesions that
appear different from erythema migrans, like African Tick Bite Fever or Rickettsiosis.

3 Conclusions and Future Work

The results of our model are a great starting point, though we aim for our results to encourage
further research in the space of skin-lesion detection for tick-borne illness diagnostics. The accuracy
rates exhibit room for improvement before being deployed into medical use. As indicated by
Rodriguez-Morales et al., more research, more diagnostics, more surveillance with better identification
approaches are very much needed. The final dataset included 6,080 images and was trained on a
combination of architectures (ResNet 34, ResNet 50, VGG 19, and DenseNet 121). We obtained an
accuracy of 80.72% with our model trained on the DenseNet 121 architecture.

For next steps, we plan to share the dataset publicly, test out different models such as a classical
model approach using a Support Vector Machine on the image pixels. We may also consider testing
an ensemble approach, as well as training on Inceptionv4, MobileNet, and EfficientNet to observe
whether these models obtain similar accuracies.
References


A Overview of sample images by search query per language

<table>
<thead>
<tr>
<th>Language</th>
<th>Tick Bite</th>
<th>Mosquito Bite</th>
<th>Healthy Skin</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>320</td>
<td>240</td>
<td>320</td>
<td>880</td>
</tr>
<tr>
<td>Spanish</td>
<td>320</td>
<td>320</td>
<td>320</td>
<td>960</td>
</tr>
<tr>
<td>French</td>
<td>240</td>
<td>320</td>
<td>320</td>
<td>880</td>
</tr>
<tr>
<td>German</td>
<td>240</td>
<td>320</td>
<td>320</td>
<td>880</td>
</tr>
<tr>
<td>Portuguese</td>
<td>240</td>
<td>320</td>
<td>240</td>
<td>800</td>
</tr>
<tr>
<td>Mandarin</td>
<td>320</td>
<td>80</td>
<td>320</td>
<td>720</td>
</tr>
<tr>
<td>Russian</td>
<td>320</td>
<td>320</td>
<td>320</td>
<td>960</td>
</tr>
<tr>
<td>Total</td>
<td>2000</td>
<td>1920</td>
<td>2160</td>
<td>6080</td>
</tr>
</tbody>
</table>

B Overview of queries for Google Image search in seven selected languages

<table>
<thead>
<tr>
<th>Language</th>
<th>Tick Bite Query</th>
<th>Mosquito Bite Query</th>
<th>Healthy Skin Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Tick bite</td>
<td>Mosquito bite</td>
<td>Healthy skin</td>
</tr>
<tr>
<td>Spanish</td>
<td>Picadura de garrapata</td>
<td>Picadura de mosquito</td>
<td>Piel saludable</td>
</tr>
<tr>
<td>French</td>
<td>Morsure de tique</td>
<td>Pique de Moustique</td>
<td>Peau saine</td>
</tr>
<tr>
<td>German</td>
<td>Zeckenbiss</td>
<td>Mückenstich</td>
<td>Gesunde haut</td>
</tr>
<tr>
<td>Portuguese</td>
<td>Mordida de carrapato</td>
<td>Picada de mosquito</td>
<td>Pele saudável</td>
</tr>
<tr>
<td>Mandarin</td>
<td>Pí yˇao</td>
<td>Wénchóng dˇìngyˇao</td>
<td>Jiˇànkˇáng de pˇífˇu</td>
</tr>
<tr>
<td>Russian</td>
<td>укус клеща</td>
<td>komarínnyy ukus</td>
<td>zdorovaya kozha</td>
</tr>
</tbody>
</table>

C Classification Results over multiple neural network architectures

<table>
<thead>
<tr>
<th>Model</th>
<th>Training Loss</th>
<th>Validation Loss</th>
<th>Validation Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ResNet-34</td>
<td>0.400625</td>
<td>0.702235</td>
<td>79.16%</td>
</tr>
<tr>
<td>ResNet-50</td>
<td>0.749036</td>
<td>1.147042</td>
<td>53.12%</td>
</tr>
<tr>
<td>VGG-19</td>
<td>0.452295</td>
<td>0.519739</td>
<td>78.64%</td>
</tr>
<tr>
<td>DenseNet-121</td>
<td>0.339066</td>
<td>0.568043</td>
<td>80.72%</td>
</tr>
</tbody>
</table>

*Table 1: Classification Results over multiple neural network architectures*
D  Results of search queries run on VGG 19 architecture

Figure 1: Results of search queries run on VGG 19 architecture in (a) English, (b) Spanish, (c) French, (d) German (e) Mandarin (f) Portuguese and (g) Russian.

E  Additional Dataset Images

Figure 2: Results from search queries in English
Figure 3: Results from search queries in Spanish

Figure 4: Results from search queries in French

Figure 5: Results from search queries in German
Figure 6: Results from search queries in Mandarin

Figure 7: Results from search queries in Portuguese

Figure 8: Results from search queries in Russian