A Personalized Diagnostic Tool for Microbiome-Related Morbidities

Olympia Giannou1*, Ahmad Mustafa Shiri2, Anastasios D. Giannou2,3, Dimitra E. Zazara4,5, Samuel Huber2#, Georgios Pavlidis1#

1 Computer Engineering & Informatics Dept., Polytechnic School, University of Patras, 26500, Patras, Greece.
2 Department of Medicine, University Medical Center Hamburg-Eppendorf (UKE), 20246, Hamburg, Germany.
3 Department of General, Visceral and Thoracic Surgery, UKE, Hamburg, Germany.
4 Division for Experimental Feto-Maternal Medicine, Department of Obstetrics and Prenatal Medicine, UKE, Hamburg, Germany.
5 Department of Pediatrics, UKE, Hamburg, Germany.

* Corresponding author. Tel.: +30-6970814929; email: ogiannou@ceid.upatras.gr
# co-senior authors
Manuscript submitted November 15, 2021; accepted February 22, 2022.

Abstract: A model-driven approach suitable for classifying microbiome-related morbidities such as ulcerative colitis on smart mobile devices is investigated in this manuscript. A novel scheme is proposed, which consists of a pre-trained image classifier on ImageNet and is deployed into the presented Android mobile application for this purpose. Endoscopic images of mouse colitis were used as input datasets for our experiments. The proposed approach offers an efficient classifier, based on the average of all its performance metrics: confusion matrix, accuracy, recall, precision, cross entropy, f1-score. The results are compared with those of the most representative image classifiers for the kind of classification we target, in terms of performance, as well as the size of the retrained frozen graph on our dataset. Such a classification could serve as a valuable tool in clinical medicine offering an automated, diagnostic tool for microbiome-related morbidities, thus allowing accurate early diagnosis and the design of personalized and targeted therapeutic approaches.

Key words: Classification, CNN, machine learning, microbiome, morbidity, diagnosis, colitis, mouse.

1. Introduction

Machine Learning (ML) and Convolutional Neural Networks (CNN) perform automatic feature extraction and pattern recognition, based on existing prior knowledge. These methods provide a simple and efficient approach, compared to the traditional image classification methods [1]. In the method of Transfer Learning (TL) [2], the CNNs, pre-trained on different data types, preserve the learned features in network structures and weights by freezing a tensorflow graph (protocol buffer .pb) [3]. The last layer is trained on the new dataset, building an accurate classifier, using much less data, without High Performance Computing (HPC), requirements.

TL is widely used for CNNs to classify images with high accuracy and achieve high pattern recognition, also in medicine. Such examples include the pre-trained architecture InceptionV3 on the ImageNet dataset [2], which achieves 70.1% accuracy on the CIFAR dataset. Skin cancer detection and classification of images with
skin cells is done with 99.77% accuracy in [4]. Liver cancer trait classification on smart mobile devices is
done with accuracy of ~82% [5]. In [6], the classification of gastric infections achieves 90% prediction
accuracy. Classification of colorectal diseases is achieved in [7] with a F1-score of 0.93 for the colorectal
dataset and 0.88 for KVASIR dataset. Classification of ulcerative colitis severity is elaborated in [8] using
endoscopic videos as input dataset.

Ulcerative colitis (UC) is an Inflammatory bowel disease (IBD), which causes irritation, inflammation and
ulceration in the lining of the large intestine (colon). These disease notes are increasing worldwide [9].
White blood cells attack the lining of the colon and cause the inflammation and ulcers. The risk for suffering from
ulcerative colitis and other morbidities is affected by age, genetics, environmental factors and changes in normal
gut bacteria [10], [11]. Importantly, changes in the microbiome have been associated with the manifestation of UC,
as well as other morbidities [12], [13]. Specifically, the host microbiome may be influenced by diet and medication,
and may even be subject to diurnal oscillation, all leading to changes that may in turn interfere with cytokine
expression patterns [14]-[17]. Cytokine expression has been shown to be dysregulated by specific factors
attributed to the microbiome and associated by-products. Under these circumstances, these factors may
contribute to the development of chronic inflammation, autoimmunity and disease severity by enhancing the
expression of specific cytokines [18]. Targeting the microbiome represents a therapeutic approach for these
inflammatory diseases [19].

In this work, our goal is to develop an efficient and accurate personalized diagnostic tool for microbiome-
related morbidities such as ulcerative colitis. An image classification of colitis, according to the detected type
of microbiome (J or Y) and the grade of colitis (severity) is developed. The input dataset consists of colon
endoscopy images of mice. The proposed classifier is targeted to smart mobile devices that report the
predicted label and confidence value for each predicted colitis microbiome category. In this paper, the TL
approach and the open source image classifier models MobileNetV1 and InceptionV3 from Google [19], pre-
machines, are used. The performance of the MV1-LCCP scheme, proposed in [5], is compared with both
MobileNetV1 and InceptionV3 on our dataset of mouse colitis images. The structure of this paper is as follows.
The material and the proposed model-driven approach are presented in Section 2. Results and comparisons
are given in Section 3 and conclusions are discussed in Section 4.

2. Materials and Proposed Scheme

2.1. Dataset—Data Collection

The following mouse experiments were performed to obtain the dataset, used in this study. Firstly, fecal
transplantation was performed to engraft the mice with different microbiome (J and Y microbiome).
Specifically, a fresh fecal microbiota transplant (FMT) was prepared by harvesting the cecal and colon
contents from normal healthy mice with different microbiome (J and Y). The contents were then resuspended
in sterile buffer and the mice received 200 μl FMT orally using gavage needle.

After 3-4 weeks, the colitis score was evaluated by endoscopy. The mice were bred and housed under
specific pathogen-free conditions in the animal facility of the University Medical Center Hamburg-Eppendorf.
Age- and sex-matched littermates from 8 to 16 weeks of age were used. Animal experiments were performed
in accordance with the Institutional Review Board “Behörde für Soziales, Familie, Gesundheit und
Verbraucherschutz” (Hamburg, Germany).

2.2. Classification and Prediction Score

All protocol buffers (python), trained models, which were generated in this work, were deployed on our
android application “Colitis Classification” (java). The steps of the proposed scheme are in Fig. 1.
**Input:** The dataset, described in Section 2.1, was used as input in our model-driven approach and was split into 3 subsets: training (80%), validation (10%) and testing (10%). Cropping, padding, horizontal flipping and rescaling were applied to fit the pre-trained model architecture. Training results were evaluated on 50 images (batch size) at a time.

**Convolutional Base (Feature Extraction):** Feature values were extracted from the input dataset using the MV1-LCCP architecture [5]. Google’s TL approach and open-source lightweight image classifier MobilenetV1 (size 0.5 at 224 pixels input image size), pre-trained on the ImageNet dataset in very fast machines, was used along with two additional fully connected dense layers before the softmax layer, to achieve better performance and accuracy in prediction (dashed blue area in Fig. 2), while providing a reasonable size of the generated retrained graph for smart mobile devices. More specifically, MobileNetV1 converted the standard convolution to a deep convolution in depth (detailed in Fig. 1). At each of these layers, batch normalization and dropout were applied. Then, the final layer was trained on our colitis dataset and the newly trained graph, frozen as a protobuf (.pb) was generated.

**Prediction:** The proposed architecture is able to predict the type of the mouse microbiome (Y or J), as well as the grade of colitis (low: when the total score is <=3.5, high: when the total score is >3.5, on a scale from 0 to 10). In particular, the prediction of the type of microbiome is of great importance, since this feature allows an early diagnosis of microbiome-related morbidities and thus a potentially timely selection of
corresponding targeted therapeutic approaches. Classes of prediction: [A] predict whether the grade of colitis is low or high if the mice have the microbiome Y, [B] predict whether the grade of colitis is low or high if the mice have the microbiome J, [C] predict whether the grade of colitis is low or high and whether the microbiome is Y or J (combination of cases A and B).

**Deployment of the newly trained model on smart mobile devices:** The generated newly trained model of the classifier was deployed on our “Colitis Classification” application (Fig. 3). A physician can use scanned endoscopic images of colitis on a mobile device and then, the predicted microbiome and colitis grade designation, as well as the confidence score for each predicted category are displayed on the screen.

![Image Collection via Mobile Camera](image1)

![Image Classification via the model](image2)

![Prediction Result & corresponding Confidence Score](image3)

Fig. 3. The GUI of our mobile application for predicting the microbiome of the scanned image, as well as the grade of colitis.

### 3. Evaluation and Experiment Results

#### 3.1. Metrics

In this manuscript, the proposed model-driven approach, using the MV1-LCCP, is evaluated and particularly, if both the type of microbiome and the grade of colitis are detected, by calculating a variety of metrics: confusion matrix, accuracy, cross-entropy, precision, recall, f1-score [21].

#### 3.2. Experiment Steps and Methods

- Endoscopic images of mice with two different types of microbiome (J and Y) were used and the grade of colitis was recorded.
- 129 images of mice with Microbiome-J and 120 images with Microbiome-Y were fed as input in our model-driven approach (Section 2.2).

#### 3.3. Results

The most representative pre-trained open-source architectures of Inceptionv3 [20] and MobileNetv1 [20], which are suitable for mobile applications, were compared with MV1-LCCP [5] on our dataset, consisting of mouse colitis images with different types of microbiome. Their performance in predicting [A] the grade of colitis in the case that the mice have the microbiome Y in Fig. 4, [B] the grade of colitis in the case that the mice have the microbiome J in Fig. 5, [C] the grade of colitis (low or high) and the type of microbiome (Y or J) in Fig. 6, was evaluated, using a variety of metrics (Section 3.1) and is depicted below in Tables 1, 2 and 3.
Table 1. Performance Case [A]

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Image Size</th>
<th>Accuracy %</th>
<th>Cross Entropy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>.pb Size (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inception V3</td>
<td>299</td>
<td>77.8</td>
<td>0.45</td>
<td>0.87</td>
<td>0.82</td>
<td>0.84</td>
<td>72.5</td>
</tr>
<tr>
<td>MobileNetV1</td>
<td>224,0.50</td>
<td>66.7</td>
<td>9.47</td>
<td>0.76</td>
<td>0.57</td>
<td>0.65</td>
<td>17.0</td>
</tr>
<tr>
<td>MV1-LCCP</td>
<td>224,0.50</td>
<td>88.9</td>
<td>6.07</td>
<td>0.85</td>
<td>0.75</td>
<td>0.78</td>
<td>7.7</td>
</tr>
</tbody>
</table>

Table 2. Performance Case [B]

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Image Size</th>
<th>Accuracy %</th>
<th>Cross Entropy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>.pb Size (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inception V3</td>
<td>299</td>
<td>62.5</td>
<td>0.55</td>
<td>0.62</td>
<td>0.56</td>
<td>0.59</td>
<td>72.1</td>
</tr>
<tr>
<td>MobileNetV1</td>
<td>224,0.50</td>
<td>37.5</td>
<td>12.83</td>
<td>0.51</td>
<td>0.49</td>
<td>0.50</td>
<td>16.9</td>
</tr>
<tr>
<td>MV1-LCCP</td>
<td>224,0.50</td>
<td>75</td>
<td>5.81</td>
<td>0.73</td>
<td>0.52</td>
<td>0.61</td>
<td>7.7</td>
</tr>
</tbody>
</table>

Table 3. Performance Case [C]

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Image Size</th>
<th>Accuracy %</th>
<th>Cross Entropy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>.pb Size (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inception V3</td>
<td>299</td>
<td>63.2</td>
<td>0.46</td>
<td>0.61</td>
<td>0.37</td>
<td>0.46</td>
<td>77.7</td>
</tr>
<tr>
<td>MobileNetV1</td>
<td>224,0.50</td>
<td>52.6</td>
<td>2.96</td>
<td>0.57</td>
<td>0.29</td>
<td>0.38</td>
<td>21.0</td>
</tr>
<tr>
<td>MV1-LCCP</td>
<td>224,0.50</td>
<td>73.7</td>
<td>1.47</td>
<td>0.58</td>
<td>0.40</td>
<td>0.47</td>
<td>10.7</td>
</tr>
</tbody>
</table>

According to the recorded values, the performance in predicting the grade of colitis is better in the case of the mice with microbiome Y. This is explained by the fact that the colitis is more intense in this case, so that each model can predict in a higher performance on average of all the values of the metrics. The model-driven scheme MV1-LCCP is deemed a more efficient classifier in both cases [A] and [B] in total, taking into account that the goal is to acquire a small-in-size trained model, which is suitable for deployment on mobile devices, and simultaneously, capable of predicting the type of microbiome as well as the grade of the associated morbidity (colitis, in this work).

According to the recorded values, the performance in predicting the grade of colitis is better in the case of the mice with microbiome Y. This is explained by the fact that the colitis is more intense in this case, so that each model can predict in a higher performance on average of all the values of the metrics. The model-driven scheme MV1-LCCP is deemed a more efficient classifier in both cases [A] and [B] in total, taking into account that the goal is to acquire a small-in-size trained model, which is suitable for deployment on mobile devices, and simultaneously, capable of predicting the type of microbiome as well as the grade of the associated morbidity (colitis, in this work).

Fig. 4. Case [A]: Performance evaluation on our dataset of mice with the microbiome Y (129 images). Accuracy (row 1), cross-entropy (row 2) over the period in steps. Training period in orange, validation period in blue. The values are smoothened by the margin of 0.4, real values are shadowed in fainter color. X-axis: steps in steps; Y-axis: accuracy (row 1), cross-entropy (row 2). Confusion matrix for each architecture (row 3). Colitis grade: high (> 3.5) or low (<=3.5). Training steps=100, learning rate=0.01, evaluation step=2.
This means that our proposal can detect and recognize the grade of colitis with a higher accuracy and precision, even compared with InceptionV3. This is also proved by looking at the confusion matrices (dark green signifies higher probability of the predicted label to be identical with the actual one). The scale of probability in which the predicted label is identical with the actual one is from 1 to 0 (dark green=1, blue= <1 and >0, white/grey=0). At this stage, it is of great interest to evaluate the performance when attempting to predict the type of microbiome together with the corresponding grade of colitis. Prediction of the type of microbiome facilitates an early diagnosis of disease manifestation.

Fig. 5. Case [B]: Performance evaluation on our dataset with mice with the microbiome J (120 images). Accuracy (row 1), cross-entropy (row 2) over the period in steps. Training period in orange, validation period in blue. The values are smoothened by the margin of 0.4, real values are shadowed in fainter color. X-axis: steps in steps; Y-axis: accuracy (row 1), cross-entropy (row 2). Confusion matrix for each architecture (row 3). Colitis grade: high (> 3.5) or low (<=3.5). Training steps=100, learning rate=0.01, evaluation step=2.

Fig. 6. Case [C]: Performance evaluation on our dataset with mice with the microbiome Y (129 images) or the microbiome J (120 images). Accuracy (row 1), cross-entropy (row 2) over the period in steps. Training period in orange, validation period in blue. The values are smoothened by the margin of 0.72, real values are shadowed in fainter color. X-axis: steps in steps; Y-axis: accuracy (row 1), cross-entropy (row 2). Confusion matrix for each architecture (row 3). Colitis grade: high (> 3.5) or low (<=3.5). Training steps=1000, learning rate=0.01, evaluation step=10.
The results in the case [C] clearly show that, when the dissimilarities in the grade of colitis, as in the case of mice with the microbiome J are not remarkable, it is more difficult for all models of classifiers to correctly predict the grade of colitis, especially using the models InceptionV3 and MV1-LCCP. Additionally, it is difficult to distinguish the type of microbiome (J or Y) when the grade of colitis is high.

4. Conclusions

We show here that the proposed model-driven approach serves as an efficient classifier on our new dataset, consisting of images of colitis in mice. The following types of prediction are feasible: [A] prediction of the grade of colitis of mice with the microbiome Y, [B] prediction of the grade of colitis of mice with the microbiome J, and [C] the prediction of the grade of colitis in both cases [A] and [B]. The performance of all CNN models was evaluated in a variety of metrics that allow us to have a complete overview of all sites.

It is worth mentioning that the compact MV1-LCCP is able to achieve quite high accuracy, precision and good f-score, as well as to generate a trained final protocol buffer (.pb) with a suitable size for smart mobile applications compared to other models. Therefore, with the proposed model-driven approach, a scientist can obtain the grade of colitis of a mouse by scanning an image through a smart mobile device, which directly reports the corresponding type of microbiome. Future work should focus on collecting many more cases of microbiome-related morbidities in mice and humans in order to optimize MV1-LCCP or alternatives to more accurately predict disease severity (low, medium, high) and microbiome alteration with the goal of implementing such an approach in the clinical, diagnostic and research practice.

Conflict of Interest

The authors do not have any conflict of interest to declare.

Author Contributions

O.G. conceived, designed and carried out most experiments, analyzed data, and wrote the paper; A.M.S., A.D.G and D.E.Z. carried out in vivo experiments, analyzed data, and wrote parts of the paper; S.H. and G.P. conceived the idea and supervised the study, designed experiments, analyzed the data, and wrote the paper. All authors reviewed and concur with the submitted manuscript.

Acknowledgment

The authors thank the Germ-free Facility at the University Medical Center Hamburg-Eppendorf for their technical assistance. Funding: This work was supported in part by the Deutsche Forschungsgemeinschaft (SFB841 to S.H. and A.D.G.), the Ernst Jung-Stiftung Hamburg (to S.H.), the European Respiratory Society/short term fellowship (to A.D.G), the Else Kröner Memorial Stipendium (to A.D.G), the Werner Otto Stiftung (to A.D.G), the Erich und Gertrud Roggenbuck Stiftung (to A.D.G), the Hamburger Krebsgesellschaft Stiftung (to A.D.G). S.H. has an endowed Heisenberg-Professorship awarded by the Deutsche Forschungsgemeinschaft.

References


Copyright © 2022 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).
Olympia Giannou is a PhD student in the group of Prof. Georgios Pavlidis, in the Laboratory of Information systems and Artificial intelligence, at the University of Patras, in Greece, focusing on machine learning and intelligent information systems. In parallel, she works as a software engineer at Leica Geosystems AG. She holds a diploma (2012) in Computer Engineering and Informatics (Dipl.-Ing.), as well as a M.Sc. (2014) in computer science and information systems from the Computer Engineering & Informatics Dept., University of Patras, in Greece. She has contributed to various research and industrial projects in information system development, embedded systems, smart daily-life applications, using a variety of tools and software methods. She has also worked as a Research Assistant in NTUA and EPFL and as a Software Engineer in Leica Geosystems, Interamerican (ACHMEA) and Ithetis.

Ahmad Mustafa Shiri was born in Kabul (Afghanistan). He completed his studies in biochemistry at the University of Hamburg, Bernhard Nocht Institute for Tropical Medicine in 2014. Dr. Shiri earned his master of science in biochemistry and his PhD in biology from the University Medical Center Hamburg-Eppendorf (UKE), in 2017 and 2020 respectively, in Germany. Since 2021, he pursues postdoctoral research at the University Medical Center Hamburg-Eppendorf (UKE), in Germany.

Anastasios D. Giannou was born in Greece. He studied Nursing at the University of Athens and subsequently, Medicine at the University of Patras in Greece. He also did two masters of science in medicinal chemistry and biochemistry, respectively, as well as a PhD at the Medical Faculty of the University of Patras in Greece. In 2015 he moved to Hamburg and joined the group of Prof. Samuel Huber at the University Medical Center Hamburg-Eppendorf (UKE) as a postdoctoral researcher. In 2019 he started his clinical residency at the Surgical Department of the UKE. Since then, he works as a clinician scientist in the Department of General, Visceral and Thoracic Surgery and the Laboratory of Molecular Gastroenterology and Immunology in UKE, focusing on cancer, metastasis and related immune responses.

Dimitra E. Zazara was born in Greece. She studied medicine at the University of Patras in Greece. After finishing her medical studies in 2013, she did a master of science in biomedical sciences at the same University. In 2015 she moved to Hamburg and joined the group of Prof. Petra Arck at the University Medical Center Hamburg-Eppendorf (UKE). In the Laboratory for Experimental Feto-Maternal Medicine she completed an MD/PhD program as part of the “Feto-Maternal Immune Cross-Talk” network, focusing on the fetal origin of childhood asthma. As a clinician scientist at the moment, Dr. Zazara works as a physician in the UKE Children’s Hospital while in parallel pursuing her research interests focusing on the prenatal programming of lung health and disease in the Laboratory for Experimental Feto-Maternal Medicine in UKE, Hamburg.
**Samuel Huber** is the director of the Department of Medicine at the University Medical Center Hamburg-Eppendorf (UKE), Germany, since 2019. There he is also head of the Section of Molecular Gastroenterology and Immunology. He studied medicine in Ulm and Mainz and received his doctorate in 2006. He is also endowed W3 Heisenberg-Professor for Intestinal Immune Regulation awarded from the DFG. His research focus is immunoregulation, inflammatory bowel disease and cancer. He is specialist in gastroenterology and internal medicine.

**Georgios Pavlidis** is a professor of Computer Engineering and informatics Dept. at Patras University, director of Information systems and Artificial intelligence Laboratory and president of Athens Network of Collaborative Expert (ANCE) NGO. He pursued his PhD in informatics and operational research, at the University of Sofia, Bulgaria. In his early career: IBM system programmer, ICT engineer, NGO founder, head of Business Administration dept. (Patras university) etc. Since 1987, he is consultant of INTRASOFT SA, Singular SA, etc. Member of different EU experts groups, member of committees for bilateral cooperation of the Greek Government. Project Manager of variety of national, EU and World Bank projects in the fields of ICT, VET, SMEs, Institutional support, Social assistance, Human resources development etc. Extensive working experience in Central/Eastern European and Mediterranean countries. Author of numerous scientific articles and books in English and Greek. Currently, member of the board of Directors of Greek-Russian Chamber of Commerce and European Profiles S.A.