

Article

DeepFMD: Computational Analysis for Malaria Detection in Blood-Smear Images Using Deep-Learning Features

Aliyu Abubakar ^{1,2,*} , Mohammed Ajuji ² and Ibrahim Usman Yahya ²¹ Centre for Visual Computing, University of Bradford, Bradford BD7 1DP, UK² Department of Computer Science, Faculty of Science, Gombe State University, Tudun Wada 760231, Nigeria; majuji@gsu.edu.ng (M.A.); usmanitgombe@gsu.edu.ng (I.U.Y.)

* Correspondence: a.abubakar6@bradford.ac.uk; Tel.: +44-7574758316

Abstract: Malaria is one of the most infectious diseases in the world, particularly in developing continents such as Africa and Asia. Due to the high number of cases and lack of sufficient diagnostic facilities and experienced medical personnel, there is a need for advanced diagnostic procedures to complement existing methods. For this reason, this study proposes the use of machine-learning models to detect the malaria parasite in blood-smear images. Six different features—VGG16, VGG19, ResNet50, ResNet101, DenseNet121, and DenseNet201 models—were extracted. Then Decision Tree, Support Vector Machine, Naïve Bayes, and K-Nearest Neighbour classifiers were trained using these six features. Extensive performance analysis is presented in terms of precision, recall, f-1score, accuracy, and computational time. The results showed that automating the process can effectively detect the malaria parasite in blood samples with an accuracy of over 94% with less complexity than the previous approaches found in the literature.

Keywords: malaria; CNN; deep learning; SVM; classification

Citation: Abubakar, A.; Ajuji, M.; Yahya, I.U. DeepFMD: Computational Analysis for Malaria Detection in Blood-Smear Images Using Deep-Learning Features. *Appl. Syst. Innov.* **2021**, *4*, 82. <https://doi.org/10.3390/asi4040082>

Received: 9 August 2021

Accepted: 20 October 2021

Published: 25 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Malaria is a disease caused by mosquitoes through bites by female *Anopheles* mosquitoes. Different types of mosquito parasites infect humans because of these bites. These include *Plasmodium Ovale*, *Plasmodium malariae*, *Plasmodium vivax*, and *Plasmodium falciparum*. The world malaria report in 2020 [1] showed that these parasites are responsible for up to 229 million estimated cases of malaria and a total of 409,000 deaths worldwide. Children aged under 5 are the most vulnerable group, accounting for 67% of global deaths caused by malaria. Africa accounted for up to 94% of global cases, and six countries accounted for approximately half of the global deaths caused by malaria: Nigeria (23%), Burkina Faso (4%), Niger (4%), the Democratic Republic of Congo (11%), United Republic of Tanzania (5%), and Mozambique (4%). Other locations affected more by malaria include South-East Asia and the Eastern Mediterranean.

Despite all the efforts and financial resources put in place, e.g., more than USD 3 billion was accumulated for malaria control and elimination in 2019, of which a USD 900 million contribution came from malaria-endemic countries, the disease is still ravaging and threatening lives.

Light microscopy using blood films is currently the most popular technique for diagnosing malaria [2]. Diagnosing malaria under a microscope is conducted by applying a patient's blood drop on a glass slide, and subsequently immersing it in a staining solution to make the parasites visible. Thick and thin blood smears are prepared; thick smears allow the parasite to be detected more efficiently than thin smears. Thin smears, on the other hand, are not without advantages, since they allow the examiner to identify species and recognise parasite stages easily [3]. Light microscopy requires no complex tools but only human expertise which makes it cheaper, affordable, and readily available. However, the biggest disadvantages of using this technique include the requirement for extensive

training of personnel to equip them with necessary skills to become proficient slide readers. Training and employing staff is highly expensive and involves a large volume of manual work.

Therefore, this study proposes the use of a state-of-the-art automated technique to detect malaria parasites in blood-smear images. Our contributions include the design of automatic procedures using transfer learning to detect blood samples infected by malaria parasites. Secondly, several deep-learning models were used for feature extraction and four machine-learning algorithms for classification are comprehensively analysed. We also provided an in-depth analysis of the computational complexity in each case.

The rest of the paper is organised as follows: Section 2 provides excerpts of related literature, Section 3 provides a detailed description of the methodology, Section 4 provides experimental results and discussion, and Section 5 provides the conclusion.

2. Literature Review

The latest trend that has boosted the performance of many non-medical domains is the use of deep learning. Deep learning is an extension of a well-known multilayer neural network that automatically learns complex data representations, also known as features. However, deep-learning models, unlike humans, require a huge database of quality and annotated data to learn and make an effective decision on future occurrences. This, perhaps, is one of the reasons the medical domain has been unable to adopt the new technology during its early proliferation because it is harder to obtain annotated training sets and many privacy concerns arise. Interestingly, the trained deep-learning models can be used to solve problems in the different but similar applications via an approach known as transfer learning. These trained deep-learning models are better known as pre-trained models, which can be described as a model that has already learned to solve a similar problem as the intended one. Transfer learning is one of the three ways to use deep-learning techniques. Other techniques include training the deep-learning model from scratch and fine-tuning the existing deep-learning models [4]. Examples of pre-trained deep-learning models include the VGG16, VGG19, ResNet50, ResNet101, ResNet152, DenseNet121, DenseNet201, AlexNet, and Xception models.

Deep learning has been applied in the medical field to address various problems such as face recognition [5–7], effective classification of skin burns [8–12], and cancer diagnosis [13–15], as well as in financial fraud detection [16,17]. Interestingly, a similar approach was adopted recently to discriminate between blood-smear images that include the Plasmodium parasite and those that do not.

In [18], the authors proposed the use of deep-learning techniques for the diagnosis of malaria diseases. This was used basically to discriminate a blood-smear sample that contains malaria and those samples that are negative. During the investigation, two well-known approaches were used: training from scratch and transfer learning approaches using 27,578 RGB images comprising infected and uninfected samples in the ratio 1:1. All images were resized into 44×44 pixels. In the first approach, 16-layered Convolutional Neural Network (CNN) architecture was used with 6 convolution layers and 3 fully connected layers. The first and the second convolution layers had 32 filters each and they were followed by the max-pooling layer. The third and the fourth convolution layers had 64 filters each and were followed by the average pooling layer, and the fifth and sixth convolution layers had 128 and 256 filters, respectively. Lastly, the architecture of the CNN model was followed by three fully connected layers each with 256 feature neurons. The model was trained with 90% of the images and tested using the remaining 10% of the unseen image samples, and achieved 97.37% accuracy. In the second approach, a pre-trained CNN model (i.e., AlexNet) was used for feature extraction and SVM for classification. This transfer learning approach achieved a 91.99% accuracy.

In another development, a study by Huq, A., et al. [19] used 27,558 images that consisted of 13,779 normal images and 13,779 malaria-infected images obtained from the National Institute of Health. Though all the images were of variable sizes, and all have

three channels, they were all rescaled to 224×224 to match the input shape of the VGG16 pre-trained model used. In addition, as part of the data preparation before the training, the data were split into three: 70% of the data from each class was labeled as training data, 30% was for validation, and 10% was taken out of the training data as testing split. The training data further shrank to 60%. Two experiments were conducted: a standard one and adversarial pieces of training. In the standard training, original images were used, and during testing, and thereafter, some adversarial images were introduced or added into the testing images. The standard accuracy and adversarial accuracy achieved were 95.96% and 29.40% respectively, in which the study showed that the model was fooled by the perturbation introduced. In the second phase of the training (i.e., adversarial training), adversarial images were incorporated into the training dataset and the algorithm was able to learn and classify adversarial images (testing split) more effectively, achieving up to 93.38% accuracy, allowing the model to learn robustly and to stand against such perturbation. It also achieved a 95.79% accuracy on standard images.

In a similar development, research by Rajaraman et al. [3] was conducted to classify 27,558 blood-smear images in which 13,779 contained malaria parasites and 13,779 were uninfected using both deep learning from scratch and a transfer learning approach. A CNN model trained from scratch contained 3 convolution layers with a max-pooling layer after each layer, 2 FC layers, and ReLU activation. The CNN model had an input shape of $100 \times 100 \times 3$ and all convolution layers had a filter size of 3×3 . The second approach of that study used a certain specific layer which according to the authors provides strong discriminatory features. These pre-trained models included AlexNet, VGG16, ResNet50, Xception, and DenseNet121, and the corresponding feature layers were fc6, block5_conv2, res5c_branch2c, block14_sepconv1, and conv_16_x2, respectively. The results showed that the trained CNN model achieved a classification accuracy of 92.7% and for the transfer learning approach, ResNet50 achieved the best recognition accuracy of 95.9% with a sensitivity of 94.7%.

Another study by Rajaraman et al. [20] proposed the use of ensemble learning using three pre-trained models (VGG19, SqueezeNet, and InceptionResNet-V2) and a custom-trained CNN model. Features of those four models were put together before final classification to discriminate images that are infected with the malaria parasite and those that are healthy. In addition, the pre-trained models were also trained independently, and a comparison was made. The result showed that VGG19 achieved a 99.32% accuracy with a sensitivity of 99.31%, outperforming all other pre-trained models including the ensemble method, which recorded an accuracy of 99.11% with a sensitivity of 98.94%.

A study by Reddy et al. [21] used a pre-trained ResNet50 model to classify 27,558 blood-smear images for malaria detection. These images were evenly distributed into infected and non-infected images. The approach in this study was the removal of the top-most (classification) layer which was originally trained to classify 1000 classes of objects. In the new scenario, a classification layer with 2 output neurons was added to solve the binary problem, while the earlier (lower) layer of the ResNet50 was frozen allowing them only to contribute features towards training the newly added layer at the top. The new model was compiled and trained using Stochastic Gradient Descent and recorded a validation accuracy of 95.4%.

Several pre-trained models—AlexNet, VGG16, NasNetMobile, Xception, Inception and ResNet50—were also used for malaria parasite detection using blood-smear images by Sriporn et al. [22]. A total of 7000 RGB blood-smear images comprising 4500 infected and 2500 uninfected were used in the study, and data augmentation was applied by rotating images by 90, 180, and 270 degrees to increase the number of the samples while preserving the detailed information. Two activation functions (ReLU and Mish), along with three different optimizers (RMSprop, Nadam, and SGD), were tested with each model. The results showed that the Xception model with the Nadam optimizer and Mish activation function achieved the best detection accuracy of up to 98.8%.

3. Materials and Method

3.1. Dataset

In this study, we used a subset of the archived blood-smear images from Chittagong Medical Hospital, Bangladesh [3]. The dataset contains infected and uninfected erythrocyte images evenly distributed into two classes. The samples of infected and uninfected images are shown in Figure 1 below.

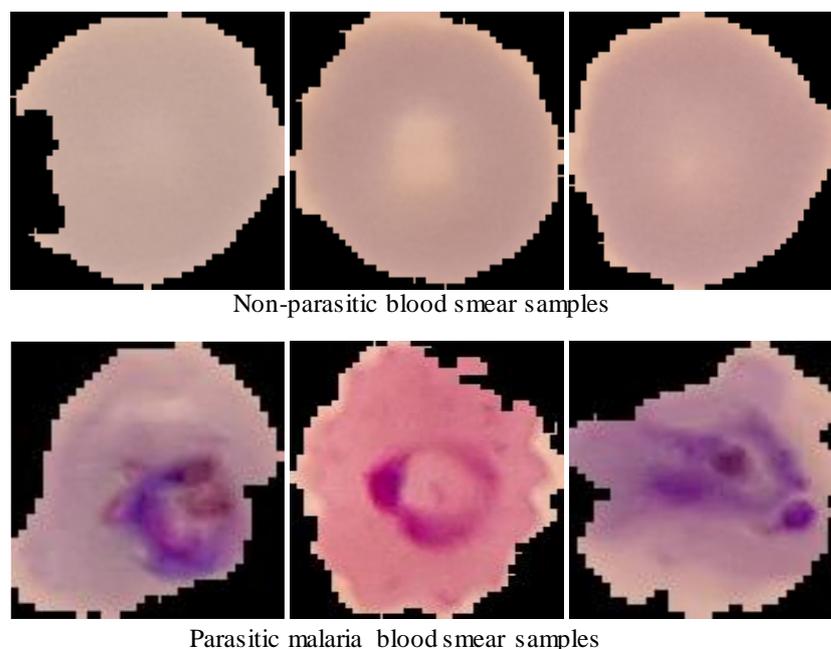


Figure 1. Sample of blood-smear images.

3.2. Dataset Pre-Processing

The images were normalised corresponding to the actual input shape of each feature extraction model, i.e., $224 \times 224 \times 3$ for VGG16, VGG19, ResNet50, and ResNet101 pre-trained CNN models. Additionally, for each DenseNet121 and DenseNet201 models, the input shape is $224 \times 224 \times 3$. All six pre-trained models were originally trained using ImageNet database [23,24], recognising and classifying 1000 distinct object classes.

3.3. Feature Extraction

For each pre-trained CNN model, top-most classification layers were removed, the remaining lower layers were retained and used only for feature extraction. Moreover, for each of the pre-trained CNN models, classification algorithms were independently used to replace the chopped top-most layers of the existing models.

3.4. Classification

Several classification algorithms can be used to classify whether a given blood-smear image contains a malaria parasite or the parasite is absent. Since we want to make a thorough comparison with six different deep-learning features, four classification algorithms were chosen based on performance and simplicity. These include Decision Tree (DT), Support Vector Machines (SVM), Naïve Bayes (NB), and K-Nearest Neighbour (KNN).

DT is used for classification problems and has an interpretability advantage over other classification algorithms [25]. It partitions data recursively into smaller subdivisions based on a set of a test at each node in the tree [26]. It contains a root node formed from the entire data set, with internal nodes or splits and the terminal nodes (also known as leaves). DT has several other advantages such as the ability to handle nonlinear relationships between feature representations, and it learns quickly [3]. SVM is one of the most popular supervised

learning algorithms used for binary classification problems. It has been applied successfully to various classification problems related to healthcare issues. [10,12]. SVM works by iteratively segregating inputs into two by finding an optimum separating hyperplane therefore producing a maximum separating distance (margin) between them. It is relatively memory-efficient and it provides clear margin distance between the classes [4]. NB is used to discriminate or classify different entities based on certain distinct features [27]. NB is based on Bayes theorem $P(A/B) = \frac{P(B/A)P(A)}{P(B)}$, where probability of A can be obtained given that B has occurred. KNN is a supervised machine-learning algorithm used for solving both classification and regression problems. It works by looking at similar features in the data presented by grouping those data points with similar features. Similar to the NB classifier, it is computationally very fast in making decisions [5,28]. The choices of selection of these classification algorithms can be attributed to their ability to work well even on small datasets and are easy to implement.

For training, 70% of the features are reserved for training and the remaining 30% for testing the trained classification algorithm. Machine-learning algorithms require lots of data samples to learn from, to generalise effectively, hence the reason for assigning a high proportion of the features for training. In addition, we applied 10-fold cross-validation on the training data split to train each model. The final prediction accuracy of the models was evaluated using the testing split. Figure 2 summarises the training process.

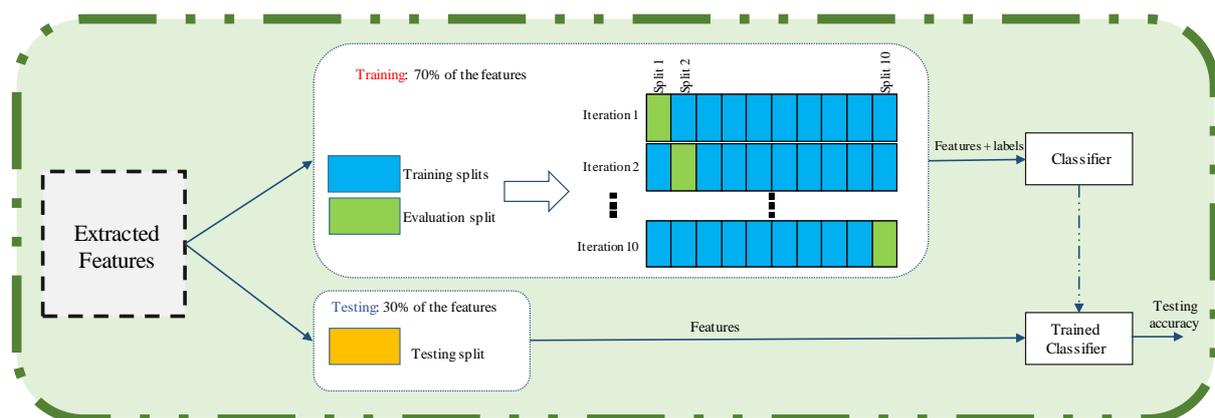


Figure 2. The feature extraction and training process.

4. Results and Discussion

Table 1 presents the results using VGG16 features. It shows the accuracy of each classification algorithm along with several performance evaluation measures, such as precision, recall (sensitivity), f1-score, and training time. Precision (also known as positive predictive value) represents a fraction of the relevant parasitic blood-smear images among the retrieved instances. Precision is mathematically represented as $\frac{TP}{TP+FP}$, where TP stands for true positive, and FP stands for false positive. Recall represents a fraction of relevant retrieved instances from parasitic blood-smear image samples. Recall is mathematically represented as $\frac{TP}{TP+FN}$, where FN stands for false negatives. F1-score is the harmonic mean of precision and recall.

Table 1. Classification results using VGG16 features.

Classification Algorithms	Performance Evaluation Metrics				
	Precision	Recall	F1-Score	Accuracy	Time (s)
DT	0.8910	0.8933	0.8922	0.8924	3528.79
SVM	0.9426	0.9557	0.9491	0.9488	71,163.34
NB	0.6283	0.9647	0.7610	0.6970	23.01
KNN	0.8866	0.9649	0.9241	0.9207	9626.79

The result in Table 1 shows the performance outputs of all four classification algorithms. These results were obtained from the VGG16 features. DT and SVM achieved precision of 89.10% and 94.26%, recall of 89.11% and 95.57%, f1-score of 89.22% and 94.91%, and accuracy of 89.24% and 94.88%, respectively. The precision for NB and KNN classifiers is 62.83% and 88.66%, recall of 96.47% and 96.49%, f1-score of 76.10% and 92.41% and accuracy of 69.70% and 92.07%, respectively.

The result in Table 2 shows performance outputs of all the four classification algorithms obtained from the VGG19 features. DT and SVM achieved precision of 86.06% and 93.84%, recall of 85.82% and 95.22%, f1-score of 85.94% and 94.52%, and accuracy of 85.94% and 94.48%, respectively. The precision for NB and KNN classifiers are 60.26% and 85.71%, recall of 96.64% and 95.15%, f1-score of 74.23% and 90.19% and accuracy of 66.46% and 89.64%, respectively.

Table 2. Classification results using VGG19 features.

Classification Algorithms	Performance Evaluation Metrics				
	Precision	Recall	F1-Score	Accuracy	Time (s)
DT	0.8606	0.8582	0.8594	0.8594	2375.60
SVM	0.9384	0.9522	0.9452	0.9448	80,246.96
NB	0.6026	0.9664	0.7423	0.6646	23.64
KNN	0.8571	0.9515	0.9019	0.8964	9635.20

The result in Table 3 shows performance outputs of all the four classification algorithms obtained from the ResNet50 features. DT and SVM achieved precision of 86.89% and 94.22%, recall of 87.17% and 95.53%, f1-score of 87.03% and 94.87%, and accuracy of 87.05% and 94.84%, respectively. The precision for NB and KNN classifiers are 80.82% and 84.69%, recall of 87.67% and 94.76%, f1-score of 84.11% and 89.44% and accuracy of 83.43% and 88.81%, respectively.

Table 3. Classification results using ResNet50 features.

Classification Algorithms	Performance Evaluation Metrics				
	Precision	Recall	F1-Score	Accuracy	Time (s)
DT	0.8689	0.8717	0.8703	0.8705	2723.95
SVM	0.9422	0.9553	0.9487	0.9484	34,231.91
NB	0.8082	0.8767	0.8411	0.8343	11.90
KNN	0.8469	0.9476	0.8944	0.8881	5029.97

The result in Table 4 shows performance outputs of all the four classification algorithms obtained from the ResNet101 features. DT and SVM achieved precision of 88.93% and 94.11%, recall of 88.40% and 95.60%, f1-score of 88.45% and 94.85%, and accuracy of 88.48% and 94.81%, respectively. The precision for NB and KNN classifiers are 82.95% and

83.11%, recall of 88.67% and 96.02%, f1-score of 85.72% and 89.10% and accuracy of 85.22% and 88.25%, respectively.

Table 4. Classification results using ResNet101 features.

Classification Algorithms	Performance Evaluation Metrics				
	Precision	Recall	F1-Score	Accuracy	Time (s)
DT	0.8893	0.8840	0.8845	0.8848	3315.88
SVM	0.9411	0.9560	0.9485	0.9481	30,682.47
NB	0.8295	0.8867	0.8572	0.8522	11.86
KNN	0.8311	0.9602	0.8910	0.8825	4988.34

The result in Table 5 shows performance outputs of all the four classification algorithms obtained from the DenseNet121 features. DT and SVM achieved precision of 87.45% and 92.55%, recall of 86.82% and 96.15%, f1-score of 87.13% and 94.32%, and accuracy of 87.08% and 94.21%, respectively. The precision for NB and KNN classifiers are 71.67% and 79.77%, recall of 73.79% and 89.27%, f1-score of 72.71% and 84.26% and accuracy of 72.31% and 83.00%, respectively.

Table 5. Classification results using DenseNet121 Features.

Classification Algorithms	Performance Evaluation Metrics				
	Precision	Recall	F1-Score	Accuracy	Time (s)
DT	0.8745	0.8682	0.8713	0.8708	1607.55
SVM	0.9255	0.9615	0.9432	0.9421	26,035.99
NB	0.7167	0.7379	0.7271	0.7231	6.91
KNN	0.7977	0.8927	0.8426	0.8300	2671.86

The result in Table 6 shows performance outputs of all the four classification algorithms obtained from the DenseNet201 features. DT and SVM achieved precision of 86.60% and 92.93%, recall of 86.14% and 95.99%, f1-score of 86.37% and 94.43%, and accuracy of 86.33% and 94.34%, respectively. The precision for NB and KNN classifiers are 68.05% and 77.95%, recall of 83.96% and 93.93%, f1-score of 75.17% and 85.20% and accuracy of 72.27% and 83.68%, respectively.

Table 6. Classification results using DenseNet201 Features.

Classification Algorithms	Performance Evaluation Metrics				
	Precision	Recall	F1-Score	Accuracy	Time (s)
DT	0.8660	0.8614	0.8637	0.8633	3612.66
SVM	0.9293	0.9599	0.9443	0.9434	50,593.37
NB	0.6805	0.8396	0.7517	0.7227	12.96
KNN	0.7795	0.9393	0.8520	0.8368	4246.80

Figure 3 depicts the comparison of each classification algorithm based on accuracy. The SVM classifier performed effectively well than the rest of the algorithms and VGG16 features carry strong discriminatory information. In terms of computational efficiency, we keep track of the training time of each classification algorithm in seconds. Tables 1–6 show the respective training times using VGG16, VGG19, ResNet50, ResNet101, DenseNet121, and DenseNet201 features, respectively.

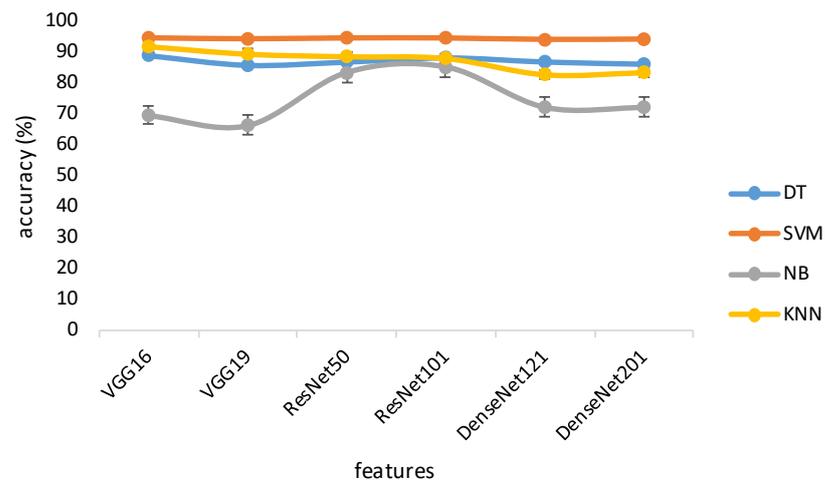


Figure 3. Comparison between the accuracy of each classifier and features.

Figure 4 shows that the SVM is computationally inefficient compared to the rest of the classifiers during the training stage, while NB is more efficient computationally than the rest of the classifiers. When we compare the NB training time, one will notice that it is the poorest classification compared to the rest of the algorithms used in this study to detect the malaria parasite in blood-smear samples. Moreover, VGG16 features provide the best detection and classification accuracy with SVM of up to 94.88% but it took more than 71163 s (19 h equivalent) to train. VGG19 features has not performed any better than VGG16 features with SVM in terms of both accuracy and computational time. Both features from VGG16 and VGG19 have 4096 feature vectors each. ResNet50 and ResNet101 have 2048 feature vectors each, and each performed better with SVM achieving classification accuracy of up to 94.84% and 94.81%, respectively, and higher computational time with ResNet50 features as seen in Figure 4. Lastly, DenseNet121 with 1024 feature vectors and DenseNet201 with feature vectors yielded the best detection accuracy of 94.21% and 94.34%, respectively along with SVM.

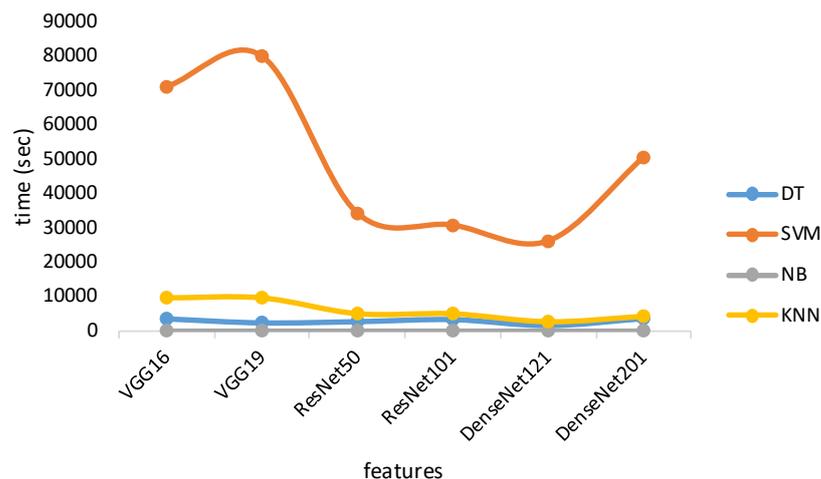


Figure 4. Computational efficiency.

One notable thing we observe is that the accuracy is higher with an increased number of feature vectors. Additionally, SVM produced outstanding results in each case, with a slight fractional difference.

Receiver Operating Characteristics (ROC) Curve

To further analyse the results obtained in this study, ROC curve is used. It has been used widely as a tool for analysing overall test performance and for the comparison of the discriminating ability of clinical tests [12,29]. It is based on the graphical curve plotting the relationship between true positive rate and false positive rate over threshold points of a test. The area under the ROC curve (AUC) summarises the overall performance estimate of the test, where AUC value below 0.5 indicates poor diagnostic test, AUC = 0.5 is considered same as random guesswork by an inexperienced clinician, AUC > 0.5 indicates a good diagnostic test, and AUC = 1 stands for perfect diagnostic test.

Figure 5 shows ROC curves of all the four classification algorithms using VGG16 features. All four classification algorithms produced impressive results, outperforming random guess with SVM achieving an almost perfect diagnostic test. The ROC curve depicted in Figure 6 was generated using VGG19 features, with SVM outperforming all the other classifiers.

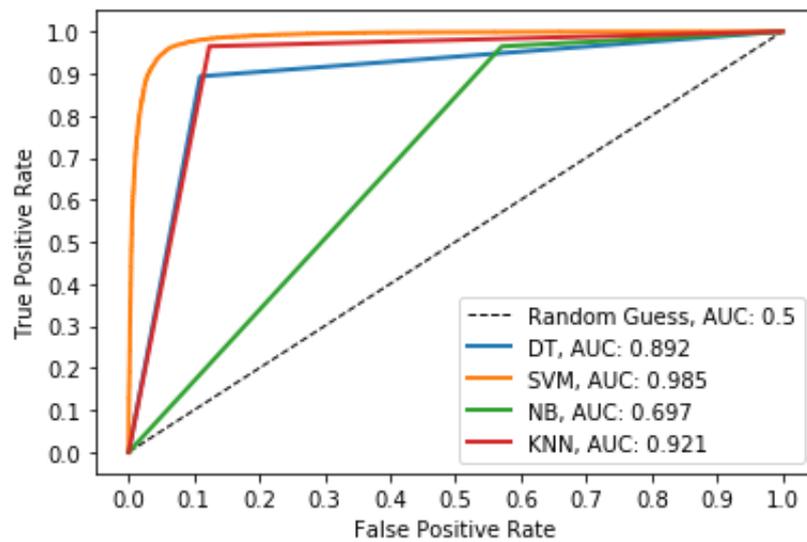


Figure 5. ROC curve and AUC using VGG16 features.

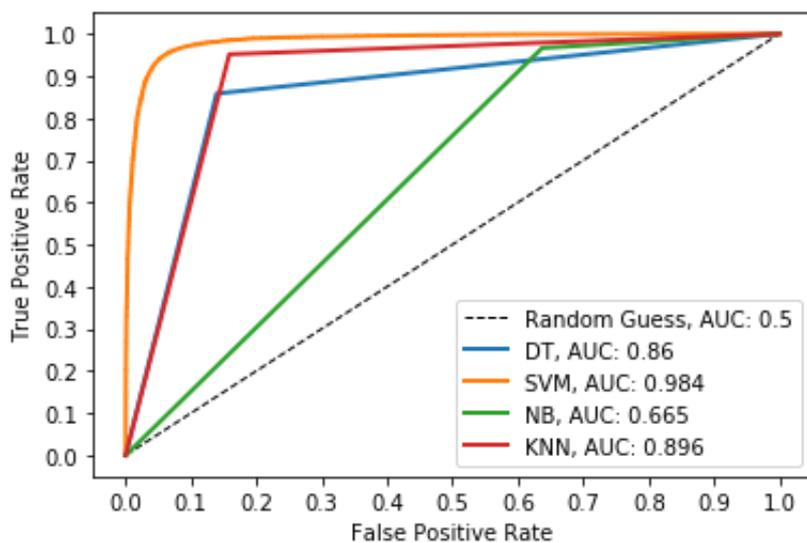


Figure 6. ROC curve and AUC using VGG19 features.

Figures 7 and 8 are the ROC curves using ResNet50 and ResNet1010 features, respectively. With ResNet50 features, SVM achieved an AUC of 0.985 and with ResNet101, the AUC is 0.987

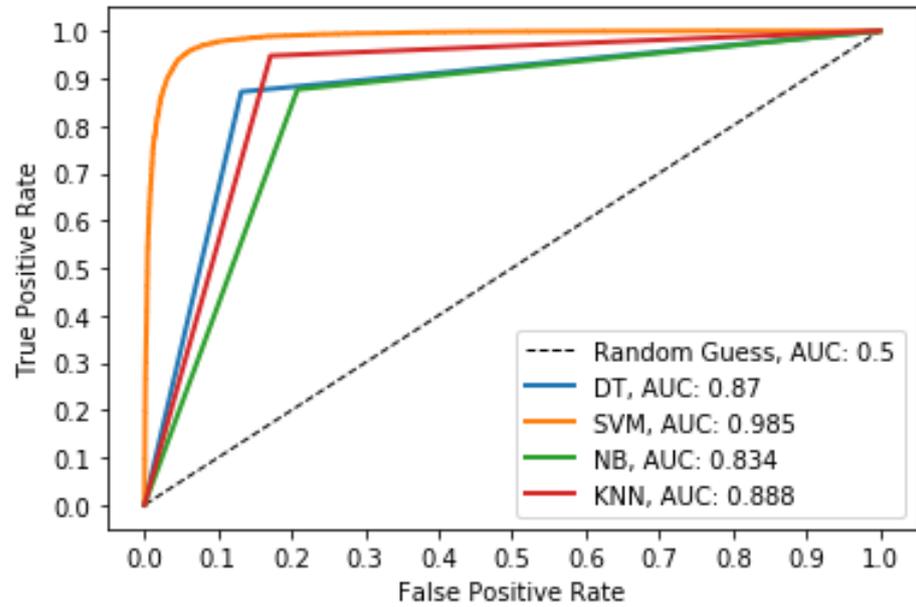


Figure 7. ROC curve and AUC using ResNet50 features.

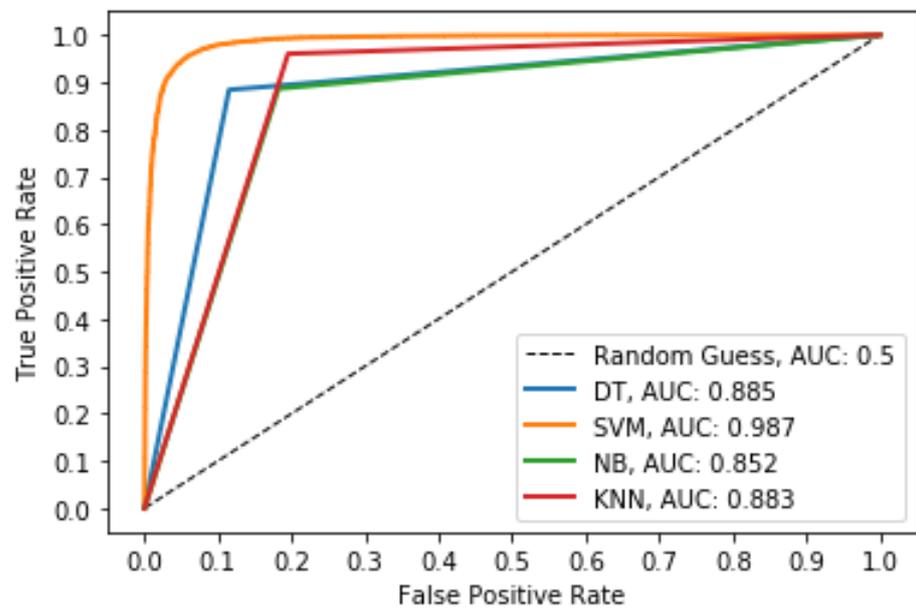


Figure 8. ROC curve and AUC using ResNet101 features.

With DenseNet121 and DenseNet201 features, both yielded 0.98 AUC score with SVM classifier as the best performing diagnostic test as shown in Figures 9 and 10, respectively

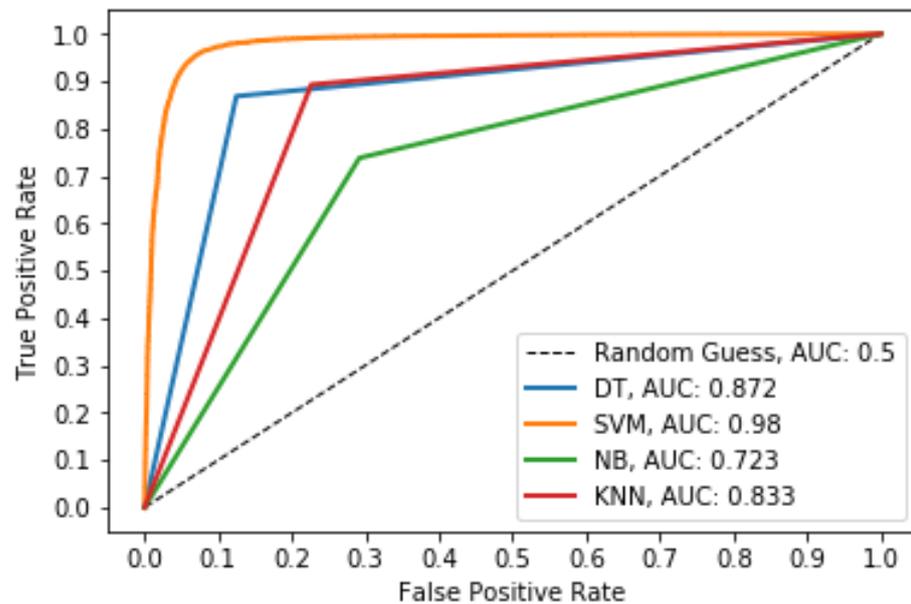


Figure 9. ROC curve and AUC using DenseNet121 features.

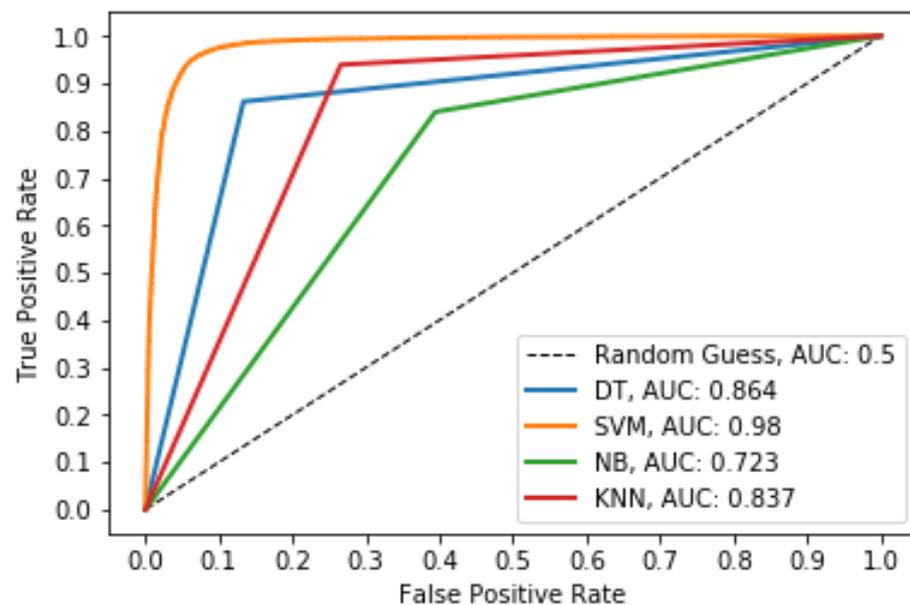


Figure 10. ROC curve and AUC using DenseNet201 features.

5. Conclusions

In this study, we presented a comprehensive investigation using state-of-the-art algorithms to effectively detect malaria parasites in blood-smear samples. We took advantage of existing pre-trained models to extract useful discriminatory features from the dataset images and subsequently used machine-learning algorithms to classify each sample to know whether a patient with a given blood sample is infected or not.

The result shows feasibility of using machine learning to detect malaria parasites in blood samples with accuracy of over 94%.

However, the dataset may contain parasites other than malaria. As such, this a limitation that is worth investigating in the future in addition to detecting each parasite (i.e., *Plasmodium Ovale*, *Plasmodium malariae*, *Plasmodium vivax*, and *Plasmodium falciparum*).

Author Contributions: Conceptualization, A.A.; methodology, A.A.; software, A.A.; validation, A.A., M.A. and I.U.Y.; formal analysis, I.U.Y.; resources, A.A.; data curation, A.A.; writing—original draft preparation, A.A.; writing—review and editing, A.A., M.A. and I.U.Y.; visualization, M.A.; project administration, A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in <https://lhncbc.nlm.nih.gov/LHC-publications/pubs/MalariaDatasets.html> (accessed on 27 June 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

- World Health Organization. *WHO Malaria Policy Advisory Group (MPAG) Meeting: Meeting Report*; April 2021; World Health Organization: Geneva, Switzerland, 2021.
- Abubakar, A.; Ugail, H.; Bukar, A.M. Assessment of human skin burns: A deep transfer learning approach. *J. Med. Biol. Eng.* **2020**, *40*, 321–333. [[CrossRef](#)]
- Lowd, D.; Davis, J. Improving Markov network structure learning using decision trees. *J. Mach. Learn. Res.* **2014**, *15*, 501–532.
- Anuradha, J.; Ramachandran, V.; Arulalan, K.; Tripathy, B. Diagnosis of ADHD Using SVM Algorithm. In Proceedings of the Third Annual ACM Bangalore Conference, Bangalore, India, 22–23 January 2010; Association for Computing Machinery: New York, NY, USA, 2010; pp. 1–4.
- Zulfikar, W.B.; Irfan, M.; Alam, C.N.; Indra, M. The Comparison of Text Mining with Naive Bayes Classifier, Nearest Neighbor, and Decision Tree to Detect Indonesian Swear Words on Twitter. In *2017 5th International Conference on Cyber and IT Service Management (CITSM), Denpasar, Indonesia, 8–10 August 2017*; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2017; pp. 1–5.
- Poostchi, M.; Silamut, K.; Maude, R.J.; Jaeger, S.; Thoma, G. Image analysis and machine learning for detecting malaria. *Transl. Res.* **2018**, *194*, 36–55. [[CrossRef](#)] [[PubMed](#)]
- Rajaraman, S.; Antani, S.K.; Poostchi, M.; Silamut, K.; Hossain, A.; Maude, R.J.; Jaeger, S.; Thoma, G.R. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ* **2018**, *6*, e4568. [[CrossRef](#)] [[PubMed](#)]
- Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, S.; Hasan, M.; Van Essen, B.C.; Awwal, A.A.S.; Asari, V.K. A state-of-the-art survey on deep learning theory and architectures. *Electronics* **2019**, *8*, 292. [[CrossRef](#)]
- Elmahmudi, A.; Ugail, H. A framework for facial age progression and regression using exemplar face templates. *Vis. Comput.* **2020**, *37*, 1–16.
- Elmahmudi, A.; Ugail, H. Experiments on Deep Face Recognition Using Partial Faces. In Proceedings of the 2018 International Conference on Cyberworlds (CW), Singapore, 3–5 October 2018; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2018; pp. 357–362.
- Jilani, S.K.; Ugail, H.; Bukar, A.M.; Logan, A. On the Ethnic Classification of Pakistani Face using Deep Learning. In Proceedings of the 2019 International Conference on Cyberworlds (CW), Kyoto, Japan, 2–4 October 2019; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2019; pp. 191–198.
- Abubakar, A. Comparative Analysis of Classification Algorithms Using CNN Transferable Features: A Case Study Using Burn Datasets from Black Africans. *Appl. Syst. Innov.* **2020**, *3*, 43. [[CrossRef](#)]
- Abubakar, A.; Ugail, H.; Smith, K.M.; Bukar, A.M.; Elmahmudi, A. Burns Depth Assessment Using Deep Learning Features. *J. Med. Biol. Eng.* **2020**, *40*, 1–11. [[CrossRef](#)]
- Abubakar, A.; Ugail, H. Discrimination of Human Skin Burns Using Machine Learning. In *Intelligent Computing—Proceedings of the Computing Conference, London, UK, 16–17 July 2019*; Springer: Cham, Switzerland, 2019; pp. 641–647.
- Abubakar, A.; Ajuji, M.; Yahya, I.U. Comparison of deep transfer learning techniques in human skin burns discrimination. *Appl. Syst. Innov.* **2020**, *3*, 20. [[CrossRef](#)]
- Ugail, H.; Alzorgani, M.; Bukar, A.; Hussain, H.; Burn, C.; Sein, T.M.; Betmouni, S. A Deep Learning Approach to Tumour Identification in Fresh Frozen Tissues. In Proceedings of the 2019 13th International Conference on Software, Knowledge Information Management and Applications (SKIMA), Ukulhas, Maldives, 26–28 August 2019; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2019; pp. 1–6.
- Singh, R.; Ahmed, T.; Kumar, A.; Kumar Singh, A.; Kumar Pandey, A.; Kumar Singh, S. Imbalanced Breast Cancer Classification Using Transfer Learning. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2020**, *18*, 83–93. [[CrossRef](#)] [[PubMed](#)]
- Hassan, S.A.; Sayed, M.S.; Abdalla, M.I.; Rashwan, M.A. Breast cancer masses classification using deep convolutional neural networks and transfer learning. *Multimed. Tools Appl.* **2020**, *79*, 30735–30768. [[CrossRef](#)]
- Craja, P.; Kim, A.; Lessmann, S. Deep learning for detecting financial statement fraud. *Decis. Support Syst.* **2020**, *139*, 113421. [[CrossRef](#)]

20. Oblé, F.; Bontempi, G. Deep-Learning Domain Adaptation Techniques for Credit Cards Fraud Detection. In *Recent Advances in Big Data and Deep Learning: Proceedings of the INNS Big Data and Deep Learning Conference INNSBDDL2019, Genova, Italy, 16–18 April 2019*; Springer: Cham, Switzerland, 2019; Volume 1, p. 78.
21. Liang, Z.; Powell, A.; Ersoy, I.; Poostchi, M.; Silamut, K.; Palaniappan, K.; Guo, P.; Hossain, A.; Sameer, A.; Maude, R.J.; et al. CNN-Based Image Analysis for Malaria Diagnosis. In *Proceedings of the 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Shenzhen, China, 15–18 December 2016*; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2016; pp. 493–496.
22. Huq, A.; Pervin, M.T. Robust Deep Neural Network Model for Identification of Malaria Parasites in Cell Images. In *Proceedings of the 2020 IEEE Region 10 Symposium (TENSymp), Dhaka, Bangladesh, 5–7 June 2020*; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2020; pp. 1456–1459.
23. Rajaraman, S.; Jaeger, S.; Antani, S.K. Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. *PeerJ* **2019**, *7*, e6977. [[CrossRef](#)] [[PubMed](#)]
24. Reddy, A.S.B.; Juliet, D.S. Transfer Learning with ResNet-50 for Malaria Cell-Image Classification. In *the Proceedings of the 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 4–6 April 2019*; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2019; pp. 0945–0949.
25. Sriporn, K.; Tsai, C.-F.; Tsai, C.-E.; Wang, P. Analyzing Malaria Disease Using Effective Deep Learning Approach. *Diagnostics* **2020**, *10*, 744. [[CrossRef](#)]
26. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]
27. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In *the Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009*; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2009; pp. 248–255.
28. Myles, A.J.; Feudale, R.N.; Liu, Y.; Woody, N.A.; Brown, S.D. An introduction to decision tree modeling. *J. Chemom.* **2004**, *18*, 275–285. [[CrossRef](#)]
29. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409. [[CrossRef](#)]