A Hybrid Approach to Histopathological Diagnosis of Paratuberculosis in Sheep with Image Classification and Explainable Artificial Intelligence with Grad-CAM

Nilgün Şengöz (nilgunsengoz@gmail.com)
Suleyman Demirel University: Suleyman Demirel Universitesi https://orcid.org/0000-0001-5651-8173
Tuncay Yiğit
Suleyman Demirel University: Suleyman Demirel Universitesi
Özlem Özmen
Burdur Mehmet Akif Ersoy Üniversitesi: Burdur Mehmet Akif Ersoy Universitesi
Ali Hakan Işık
Burdur Mehmet Akif Ersoy Üniversitesi: Burdur Mehmet Akif Ersoy Universitesi
Jude Hemanth
Karunya Institute of technology and Science: Karunya Institute of Technology and Sciences

Research Article

Keywords: paratuberculosis, deep learning, histopathology, image classification

Posted Date: February 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1318739/v1

Abstract

Paratuberculosis is a chronic, progressive, and economically important disease. The diagnosis of the disease, especially in the early stage, is difficult. Deep learning is an emerging method in the histopathological diagnosis of images. Deep Learning is the system that can perform automatic feature extraction in training sets containing tagged images and that enables the use of multi-layered neural networks, unlike machine learning. In this article, deep learning methods were used for the diagnosis of paratuberculosis in sheep's gut images. Normal intestinal histological images were used as control data. This study revealed that the deep learning method can be used with high accuracy for the histopathological diagnosis of paratuberculosis. According to Deep Learning models, the highest accuracy is proposed model Fastai/VGG-16 which has Classification Accuracy 98,36 not only this one but also the Matthews Correlation Coefficient (MCC) is 0,967. According to classification accuracy, Gradient-weighted Class Activation Mapping (Grad-CAM) was used to better understand the explainability and safety level of the model. The aim of this study, build a model that was developed to help pathologists diagnose paratuberculosis disease by transferring the diagnostic information of pathologists to deep learning models and also present an explainable model for paratuberculosis, an image-based diagnostic problem in which deep learning processes can be applied effectively.

1. Introduction

Paratuberculosis, also known as Johne's disease, is a systemic, chronic progressive, and often serious disease of ruminants. The disease is caused by an acid-fast organism *Mycobacterium avium subsp. paratuberculosis* (Brown et al., 2007; Castellanos et al., 2012; Windsor, 2014). Clinical signs generally include watery diarrhea and weight loss. But the disease may be insidious in small ruminants. Affected animals characteristically exhibit progressive weight loss, soft faces, and exercise intolerance (Windsor, 2014). The effects of the subclinical disease are less defined but may include decreased milk production; they also cause large economic losses in the livestock industry (Juste and Casal, 1993; Hutchinson, 1996).

Deep Learning, a subset of machine learning, is an artificial intelligence method that uses multilayered artificial neural networks (ANN) in field such as object detection, speech recognition, natural language processing (NLP) etc. Unlike traditional machine learning methods that use coded rules, deep learning methods enable learning automatically from the symbols of data belonging to pictures, videos, sounds, and texts. Since they are flexible, they can also learn from raw images or text data, and their estimation accuracy can increase depending on the size of the data. However, deep learning carries out the learning process through examples (Kaya, 2019).

Using deep learning models means using black box models. In such studies with black box models, people/doctors do not understand the limitations of the models and cannot answer the question of how they achieve good diagnostic results. It has been determined that it is necessary to use additional solutions to support the interpretability of traditional machine learning models or to make deep learning

models explainable. To produce 'visual explanations' for decisions made from Convolutional Neural Network (CNN) based models, a technique is essential that makes them more transparent and explainable.

While Deep Learning models provide superior performance, the inability to decompose them into individual heuristic components makes them difficult to interpret. As a result, when today's smart systems fail, they often fail without warning or explanation, leaving the user looking at inconsistent output and wondering why the system is doing what. This is why interpretability is important. Clearly, to build trust in intelligent systems and move towards their meaningful integration into our daily lives, we need to build 'transparent' models capable of explaining what they predict and why. Generally speaking, this transparency and ability to explain is useful at every stage of Artificial Intelligence (AI) evolution. Intuitive and graphical indicators of how CNN makes decisions help users to have confidence in the model.

Class activation maps are a technique for obtaining distinctive image regions used to identify a particular class in an image. In other words, you will be able to see which parts your trained model focuses on when classifying the image. Compared to other CAM Models, the distinctive regions highlighted by the Grad-CAM procedure are always focused where they are needed in the picture and never the background of the images. The basic idea behind Grad-CAM is to take advantage of the spatial information preserved through convolution layers to understand which parts of an input image are important for a classification decision.

The following set of methods was developed based on convolutional neural networks or a combination of CNN with handcrafted features. As the training of convolutional neural network models is often complex and requires a more extensive training set, initial studies (Kashif et al., 2016; Xing et al., 2016; Romo-Bucheli et al., 2016; Wang et al., 2016) considered CNN to be biologically interpretable focused on integrating construction with features, and these models have shown perfect performance in tackling the touch core segmentation problem.

Deep learning (DL) methods have also been studied to normalize histology images. Color normalization is an important field of research in histopathological image analysis. Janowczyk et al. (2016) presented a method for the stain normalization of histopathology images stained with H&E based on deep sparse autoencoders.

Sethi et al. (2016) emphasized the importance of color normalization for CNN-based tissue classification in H&E dyed images. Proposing a hybrid method based on persistent homology, Qaiser et al. (2019) were able to capture the degree of spatial connectivity between the touching cores, which is quite difficult to obtain using CNN models (Sabour et al., 2017).

For a problem that the machine is asked to solve, it is sufficient to give a model that enables the machine to find a solution to the problem by evaluating the examples instead of using rule sets. In order to correct the error in the solution of the problem, a simple command list is given, and the machine is expected to perform the learning process. Model selection is effective in solving the problem. The model to be

determined in accordance with the problem contributes more to the solution of the problem (Buduma, 2015).

As a result of the literature review, it has been seen that there is no similar study using deep learning to diagnose paratuberculosis. Digital imaging, that is, an auxiliary diagnostic tool, has not been developed for the diagnosis of paratuberculosis. Hence, to the best of our knowledge, the present study is the first study in this field. We tried many deep learning models for this study, and we present the models with the highest performance.

The aim of this study is to present an approach to histopathological diagnosis of paratuberculosis in sheep with image classification and explainable artificial intelligence with grad-cam model.

2. Deep Learning Algorithms

Artificial Intelligence (AI) is generally defined as the ability of computers to transfer the working structure of human intelligence to computers in order to perform tasks that require logic, such as drawing conclusions, finding solutions, making generalizations, understanding the problem, and learning by making use of past experiences. Artificial intelligence is a science that tries to copy or imitate human intelligence in a computer system. Deep learning (DL) is a machine learning (ML) technique that is recognized through artificial neural networks and inspired by the working principle of brain neurons. Unlike task-specific algorithms, deep learning is a machine learning method based on data representations and provides advantages in many applications because it can learn from large amounts of unstructured or unlabeled data (Figure 1). Compared to other machine learning approaches, DL has the advantage of building deep architectures to learn more abstract from data to get information. The most important feature of DL methods is that they can learn feature representations automatically, thus avoiding too many time-consuming operations.

In this section, the deep learning architectures mentioned in the article will be explained in detail.

2.1. Convolutional Neural Network (CNN)

CNN architecture, which is used extensively for image recognition problems, was first introduced in (Fukushima, 1980), and its use was first suggested in (LeCun et al., 1998). CNN architecture generally includes convolution and jointing layers. Feature maps are created with the convolution layer. Different filters can be applied in each convolution layer (Coşkun et al., 2017).

Mathematically, each filtering operation performed by a feature map is a separate convolution. As a result of the filtering process that takes place in the convolution layers, the commoning layers combine semantically similar features again. After the last jointing process, the image is flattened and classified according to the detected features (Lecun et al., 2015) (Figure 2).

2.2. ResNet

Microsoft ResNet architecture is formed from a block (Residual Block) that is fed once a residual value (Residual Value) between two RELUs and linear layers. With this structure, it was thought that learning would take place faster. Figure 4 shows this residual modulus (He et al., 2016).

Microsoft Resnet was the winner of the ILSVRC ImageNet competition held in 2015. An error rate of 3.6% in this competition was achieved. While humans classify images with an average error rate of 5-10%, an error rate of 3.6% means better visual recognition than humans. The number of layers of this architecture, which has a depth above the previous architectures, higher than the number of layers in other deep learning architectures (Russakovsky, 2015).

2.3. VGG16

Another CNN architecture used for classification is VGG-16 architecture that consists of fully connected, convolutional layers. It consists of 21 main layers in total (Simonyan et al., 2014). This architecture has an incremental network structure. The image input resolution is 224×224 pixels. The filter size in the convolutional layer is 2×2 or 3×3 pixels. In this architecture, the last layers consist of fully connected layers used for feature extraction. In this architecture, there are three cascading fully connected (FC-FullConnected) layers. There are 1,000 neurons in the last fully connected layer, and the classification layer to be produced for output has a soft-connected (SoftMax) layer (Figure 4).

3. Experimental Results And Discussion

3.1. Histopathological examination:

During the necropsy, gut samples, collected from 20 control sheep and 20 sheep with natural paratuberculosis, were fixed in 10% buffered formalin. After two-day fixation, samples were routinely processed by an automatic tissue processor (Leica ASP300S, Wetzlar, Germany) and embedded in paraffin wax. Sections of 5–µm thickness were taken from the paraffin blocks with a rotary microtome (Leica RM2155, Leica Microsystems, Wetzlar, Germany). Then, the sections were stained with hematoxylin-eosin (HE), mounted with a coverslip, and examined under a light microscope. The diagnosis was confirmed by the Zielh-Neelsen method for mycobacteria.

3.2. Data

The original data set used in this article was provided in JPG format by Burdur Mehmet Akif Ersoy University (MAKU), Faculty of Veterinary Medicine, Department of Pathology. The data set consisted of a total of 520 high-resolution (1920x1200) histology microscopic images. These images were labeled in two classes: patients with paratuberculosis and healthy intestines. There were 260 data from each class (Figure 5).

3.3. Methodology

As can be inferred from Figure 6, there were five (5) phases that had to be implemented. After the original dataset creation, we allocated it into two as test and train data. 80% of the data was reserved for the training set and the rest for the testing set. In order for the deep learning algorithm to work more effectively, the histopathology images in the dataset were preprocessed.

Histogram equalization (HE) is a common method used to increase the contrast of the image. Although HE has high speed, it often causes excessive boost and loss of local information. Adaptive histogram equalization (AHE), on the other hand, can improve local contrast by calculating several histograms corresponding to different parts of the image. Nevertheless, it has disadvantages such as long computation time and noise growth. CLAHE is the bounded model of AHE. CLAHE prevents growth by clipping the histogram at the pre-given value. It provides the advantage of not discarding the part that exceeds the clipped value and evenly distributing it on the histogram. In this context, data augmentation was also accomplished.

In this article, as mentioned above, three (3) kinds of DNN architecture were applied to the dataset to obtain the best one. F1 score and test and train accuracy values were used to compare the convolutional neural network models used.

In order to make the images more transparent and explainable, we include the trained weights from the proposed VGG16 model into the Class Activation Mapping method. It is called Gradient-weighted Class Activation Mapping (Grad-CAM). Grad-CAM uses feature maps produced by the final convolution layer of a CNN algorithm. The final convolutional layers are expected to have the best explanation between high-level semantics and detailed spatial information. As can be seen in figure 8; With the proposed Fastai/VGG16 model, it is shown where it mainly points in the classification of images. This situation, which was also approved by a specialist pathologist, shows how well the model works.

There were no pathological findings observed in the control sheep gut. A marked increase in the thickness of the gut wall was observed in paratuberculosis cases. In addition, a characteristic chronic granulomatous reaction was observed in propria mucosa. Inflammatory cells were mainly composed of lymphocytes, macrophages, epithelioid cells, and Langhans giant cells. Only Zielh-Neelsen positive cases were included in the paratuberculosis cases.

The efficiency of the classifiers in dividing test images into two classes is measured against three performance measures, such as classification accuracy, sensitivity, and specificity. These performance measures are estimated for each category, and then average measures of the two categories are determined to evaluate the performance of the classifier (Hemanth, et al, 2011). The formulas for calculating these performance measures are given as follows;

$$ClassificationAccuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

1

$$Sensitivity(Recall) = \frac{TP}{TP + FN}$$

2

$$Specificity = \frac{TN}{TN + FP}$$

3

$$Precision = \frac{TP}{TP + FP}$$

4

$$F1\,Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

5

In where TP corresponds to True Positive, TN corresponds to True Negative, FP corresponds to False Positive and FN corresponds to False Negative. TP and TN; return the number from which the classes are guessed correctly. FP and FN; return the number of incorrect predictions of classes with each other.

Confusion matrix was used to interpret the results of an established classification model and to examine the errors in the relationship between the actual and predicted values crosswise. Positive and Negative terms in the matrix do not represent correctness or inaccuracy, but rather the classes to be separated.

Another part of the evaluate binary classification is Matthews Correlation Coefficient (MCC). Accuracy is a metric that is widely used to measure the success of a model but does not appear to be sufficient on its own. An alternative metric not affected by the problem of unbalanced datasets; the MCC is a contingency matrix method for calculating the Pearson product moment correlation coefficient between actual and estimated values. MCC formula shown in below;

$$MCC = \frac{TP * TN - FP - FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}}$$

6

After applying the values obtained from the confusion matrix to the MCC formulation, the worst value is -1, and the best value is +1.

Table 1
Confusion Matrix of Own Model

Classification Accuracy	Sensitivity	Specificity	Precision	F1 Score	MCC	Training Time (s)
0,956284153	0,967	0,946	0,945	0,95604	0,913	720

Table 2
Confusion Matrix of Fastai/ResNet50

Classification Accuracy	Sensitivity	Specificity	Precision	F1 Score	MCC	Training Time (s)
0,972677596	0,968	0,978	0,978	0,97297	0,945	1500

Table 3
Confusion Matrix of Proposed Fastai/VGG-16

Classification Accuracy	Sensitivity	Specificity	Precision	F1 Score	MCC	Training Time (s)
0,983606557	0,978	0,989	0,989	0,98360	0,967	585

3.4. Discussion

The present study was undertaken to use deep learning methods for the diagnosis of paratuberculosis in sheep. The characteristic lesions in this disease are mainly located in the intestines and mediastinal lymph nodes. Although typical lesions occur in the intestine in chronic cases, lesions may be slight and confusing in new cases. Sometimes, inexperienced veterinarians may make misdiagnoses because of slight lesions. Recently, the use of deep learning methods for pathological diagnosis has been increasing rapidly. Most of the research is about human diseases, especially tumor diagnosis. This study used deep learning methods for the diagnosis of paratuberculosis in sheep and achieved a high level of accuracy.

The increasing data of large scale of full-slide images (WSI) of tissue samples has made digital pathology and also microscopy popular application study field for ANN techniques. Improved techniques applied to this area focus on three broad areas. The difficulties are divided into (1) detection, segmentation, or classification of nuclei, (2) segmentation of major organs, and (3) detection and classification of the disease of interest at the lesion or WSI level. In this study, both chronic and newly formed slight cases were histopathologically evaluated.

The use of deep learning systems for pathological diagnosis is a very new technique. Nevertheless, it is rapidly becoming widespread because of the high ratio of true and fast diagnoses. This study showed that deep learning can be used for the diagnosis of paratuberculosis.

3.5. Future Works

Within the scope of this study, deep learning applications has been modeled to be used for histopathologic images and the security level of the said model has been put into use in an explainable way thanks to the Grad-CAM technique.

The problem of being a black-box, especially seen by deep learning techniques, which is used in solving an important problem such as paratuberculosis, to be transformed into an understandable structure for doctors, veterinarians etc. with extra efforts. Although there are alternative techniques that enable the 'explainability' factor in order to ensure that such models are understandable and safe, it becomes quite easy to determine which regions of the image-based input data are used in the decision process through the heat maps obtained by the use of Grad-CAM. In this context, the Fastai\VGG16-Grad-CAM system, which was established by the integration of Grad-CAM into the Convolutional Neural Network in the aforementioned study, provided an explainable approach in the detection of paratuberculosis and ensured a healthy interaction between the human and the system.

The data set to be shared with the employees involved in this subject can allow the development of different models. In addition, a mobile application can be developed for this model to be used more quickly and easily by expert pathologists. Thus, veterinarians who do not have sufficient knowledge on this subject can benefit from the model.

In the study, deep learning models were developed after putting the data in the preprocessing stage with the CLAHE method. In this context, new pretreatment methods can be applied in future studies, and comparisons can be made based on this article.

Declarations

Ethical approval Ethical approval was not required as all the dataset used in this study were animals that had died due to the disease mentioned in the article.

Funding details This study was not funded by any institution.

Conflict of interest The authors of the manuscript declare that they haveno affiliations with or involvement in any organization or entity withany financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies,stock ownership, or other equity interest; and expert testimony or patentlicensing arrangements), or non-financial interest (such as personal orprofessional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Nilgün Şengöz, Özlem Özmen,TuncayYiğit, Jude Hemanth and Ali HakanIşık. The first draft of the manuscript was written by Nilgün Şengöz and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript

References

1. Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S., Navab, N., 2016. Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. IEEE Transactions on Medical Imaging 35, 1313–1321
2. Amgad, M., Elfandy, H., Hussein, H., Atteya, L.A., Elsebaie, M.A., Abo Elnasr, L.S., Sakr, R.A., Salem, H.S., Ismail, A.F., Saad, A.M., et al., 2019. Structured crowdsourcing enables convolutional segmentation of histology images. Bioinformatics 35, 3461–3467
3. Brooks BW, Robertson RH, Corner AH, Samagh BS, Garcia MM, Turcotte C, et al., 1988, Evaluation of the serological response of sheep in one flock to Mycobacterium paratuberculosis by crossed immunoelectrophoresis, Can J Vet Res, 52(2): 199
4. Brown CC, Baker DC, Barker IK, 2007, Alimentary System, In: Maxie MG, editor, Jubb, Kennedy, and Palmer's Pathology of Domestic Animals, London: Elsevier Saunders, 222-225
5. Brown CC, Baker DC, Barker IK, 2007, Alimentary System, In: Maxie MG, editor, Jubb, Kennedy, and Palmer's Pathology of Domestic Animals, London: Elsevier Saunders, 222-225.
6. Buduma, N. (2015). Fundamentals of deep learning, Copyright © 2015 Nikhil Buduma. All rights reserved. Printed in the United States of America. Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. November 2015, First edition
7. Buergelt CD, Hall C, Mc Entee K, Duncan JR, 1978, Pathological Evaluation of Paratuberculosis in Naturally Infected Cattle, Vet Pathol, 15: 196-207
8. Burnside DM, Rowley BO, 1994, Evaluation of an enzyme-linked immunosorbent assay for diagnosis of paratuberculosis in goats, Am J Vet Res, 55(4): 465
9. Carrigan MJ, Seaman JT, 1990, The pathology of Johne's disease in sheep. Aust Vet J, 67(2):47-50.
10. Castellanos E, De Juan L, Dominguez L, Aranaz A, 2012, Progress in molecular typing of Mycobacterium avium subspecies paratuberculosis, Res Vet Sci, 92(2): 169-179
11. Cireşan, D.C., Giusti, A., Gambardella, L.M., Schmidhuber, J., 2013. Mitosis detection in breast cancer histology images with deep neural networks, in: International Conference on Medical Image Computing and Computer-assisted Intervention, pp. 411–418
12. Clark, A.; Contributors. Python Imaging Library (Pillow fork). 2011, https://github.com/python-pillow/Pillow.
13. Clarke CJ, 1997, The pathology and pathogenesis of paratuberculosis in ruminants and other species. J Comp Path 116,217–261.
14. Clarke CJ, Little D, 1996, The pathology of ovine paratuberculosis: gross and histological changes in the intestine and other tissues, J Comp Pathol, 114(4):419-437.
15. Collins DM, Hilbink F, West DM, Hosie BD, Cooke MM, De Lisle GW, 1993, Investigation of Mycobacterium paratuberculosis in sheep by faecal culture, DNA characterization and the polymerase chain reaction, Vet Rec, 133(24):599.

16. Coşkun, M., Yıldırım, Ö., Uçar, A., & Demir, Y. (2017). An Overview Of Popular Deep Learning Methods. European Journal of Technic (EJT), 165-176.
17. D. Jude Hemanth, et al., Performance Improved Iteration-Free Artificial Neural Networks for Abnormal Magnetic Resonance Brain Image Classification, Neurocomputing (2013), http://dx.doi.org/10.1016/j.neucom.2011.12.066i
18. Dubash K, Shulaw WP, Bech-Nielsen S, Stills HF Jr, Slemons RD, 1995, Evaluation of an enzyme-linked immunosorbent assay licensed by the USDA for use in cattle for diagnosis of ovine paratuberculosis, J Vet Diagn Invest, 7: 347.
19. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118.
20. Fukushima, K. (1980). Neocognitron: A Self-organizing Neural Network Model for a Mechanism. Biol. Cybernetics 36, 193-202.
21. Gao, Z., Wang, L., Zhou, L., Zhang, J., 2017. Hep-2 cell image classification with deep convolutional neural networks. IEEE Journal of Biomedical and Health Informatics 21, 416–428
22. Gao, Z., Wang, L., Zhou, L., Zhang, J., 2017. Hep-2 cell image classification with deep convolutional neural networks. IEEE Journal of Biomedical and Health Informatics 21, 416–428
23. Gezon HM, Bither HD, Gibbs HC, 1988, Identification and control of paratuberculosis in a large goat herd, Am J Vet Res, 49(11): 1817.
24. Gillan, S., O'Brien, R., Hughes, A.D., Frank, T.J., Griffin, J.F.T., 2010. Identification of immune parameters to differentiate disease states among sheep infected with Mycobacterium avium subsp. Paratuberculosis. Clin. Vaccine Immunol. 17, 108–117.
25. Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., Webster, D. R., Dec. 2016. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316, 2402–2410.
26. Gutierrez Cancela MM, Garcia Marin JF, 1993, Comparison of Ziehl–Neelsen staining and immunohistochemistry for the detection of Mycobacterium bovis in bovine and caprine tuberculous lesions, J Comp Pathol, 109, 361–370.
27. He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep residual learning for image recognition. arXiv:1512.03385
28. Hilbink F, West DM, De Lisle GW, Kittelberger L, Hosie BD, Hutton J, et al., 1994, Comparison of a complement fixation test, a gel diffusion test and two absorbed and unabsorbed ELISAs for the diagnosis of paratuberculosis in sheep, Vet Microbiol, 41: 107
29. Howard, J. et al. (2018). fastai. https://github.com/fastai/fastai
30. Howard, J.; Ruder, S. Fine-tuned Language Models for Text Classification. CoRR 2018, abs/1801.06146

31. Hutchinson LJ, 1996, Economic impact of paratuberculosis, Vet Clin North Am Food Anim Pract, 12, 373–381
32. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. CoRR 2015, abs/1502.03167
33. Janowczyk, A., Basavanhally, A., Madabhushi, A., 2016. Stain normalization using sparse autoencoders (StaNoSA): Application to digital pathology. Comput Med Imaging Graph, in press
34. Juste RA, Casal J, 1993, An economic and epidemiologic simulation of different control strategies for ovine paratuberculosis, Prev Vet Med, 15 (2-3): 101, 1993.
35. Juste RA, Marco JC, De Ocariz CS, Aduriz JJ, 1991, Comparison of different media for the isolation of small ruminant strains of Mycobacterium paratuberculosis, Vet Microbiol, 28 (4): 385-390.
36. Kashif, M.N., Raza, S.E.A., Sirinukunwattana, K., Arif, M., Rajpoot, N., 2016. Handcrafted features with convolutional neural networks for detection of tumor cells in histology images, in: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 1029– 1032
37. Kaya U., Yılmaz A. (2019). Deep Learning, 1-2, ISBN:978-605-2118-399
38. Kreeger JM 1991, Ruminant paratuberculosis – a century of progress and frustration. J Vet Diagn Invest 3,373–383.
39. Krizhevsky, A., Sutskever, I., Hinton, G., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. pp. 1097–1105
40. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE 86, 2278–2324
41. Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
42. Lin, M., Chen, Q., Yan, S., 2013. Network in network. arXiv:1312.4400
43. Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., Venugopalan, S., Timofeev, A., Nelson, P. Q., Corrado, G. S., Hipp, J. D., Peng, L., Stumpe, M. C., 2017. Detecting cancer metastases on gigapixel pathology images. arXiv:1703.02442
44. McKinney, W. Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference; van der Walt, S.; Millman, J., Eds., 2010, pp. 51 – 56
45. Milner AR, Mack WN, Coates KJ, 1989, A modified ELISA for the detection of goats infected with Mycobacterium paratuberculosis, Aust Vet J, 66(9): 305.
46. Oliphant, T. NumPy: A guide to NumPy. USA: Trelgol Publishing, 2006
47. Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic Differentiation in PyTorch. NIPS Autodiff Workshop, 2017
48. Qaiser, T., Tsang, Y.W., Taniyama, D., Sakamoto, N., Nakane, K., Epstein, D., Rajpoot, N., 2019. Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features. Medical Image Analysis 55, 1–14
49. Rao, S., 2018. Mitos-rcnn: A novel approach to mitotic figure detection in breast cancer histopathology images using region based convolutional neural networks. arXiv preprint

arXiv:1807.01788

50. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). Imagenet large scale visual recognition challenge. International Journal of Computer Vision, 115(3), 211-252.
51. Sabour, S., Frosst, N., Hinton, G.E., 2017. Dynamic routing between capsules, in: Advances in Neural Information Processing Systems, pp. 3856–3866
52. Sethi, A., Sha, L., Vahadane, A. R., Deaton, R. J., Kumar, N., Macias, V., Gann, P. H., 2016. Empirical comparison of color normalization methods for epithelial-stromal classification in H and E images. J Pathol Inform 7, 17
53. Sherman DM, Gezon HM, 1980, Comparison of agar gel immunodiffusion and fecal culture for identification of goats with paratuberculosis, J Am Vet Med Assoc, 177(12): 1208-1211
54. Shulaw WP, Bech-Nielsen S, Rings DM, Woodruff TS, 1993, Serodiagnosis of paratuberculosis in sheep by use of agar gel immunodiffusion, Am J Vet Res, 54(1): 13-19.
55. Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556
56. Sugden EA, Brooks BW, Young NM, Watson DC, Nielsen KH, Corner AH, et al., 1991, Chromatographic purification and characterization of antigens A and D from Mycobacterium paratuberculosis and their use in enzyme-linked immunosorbent assays for diagnosis of paratuberculosis in sheep, J Clin Microbiol, 29 (8): 1659-1664.
57. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2014. Going deeper with convolutions. arXiv:1409.4842
58. Valkonen, M., Isola, J., Isola, J., Ylinen, O., Muhonen, V., Saxlin, A., Tolonen, T., Nykter, M., Ruusuvuori, P., 2019. Cytokeratin-supervised deep learning for automatic recognition of epithelial cells in breast cancers stained for ER, PR, and Ki-67. IEEE Transactions on Medical Imaging , 1–1.
59. Wang, S., Yao, J., Xu, Z., Huang, J., 2016. Subtype cell detection with an accelerated deep convolution neural network, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 640–648
60. Windsor, P.A. 2015. Paratuberculosis in sheep and goats. Vet. Microbiol. 181, 161–169.
61. Xing, F., Xie, Y., Yang, L., 2016. An automatic learning-based framework for robust nucleus segmentation. IEEE Transactions on Medical Imaging 35, 550–566; Romo-Bucheli, D., Janowczyk, A., Gilmore, H., Romero, E., Madabhushi, A., 2016. Automated tubule nuclei quantification and correlation with oncotype DX risk categories in ER+ breast cancer whole slide images. Scientific Reports 6, 32706

Figures

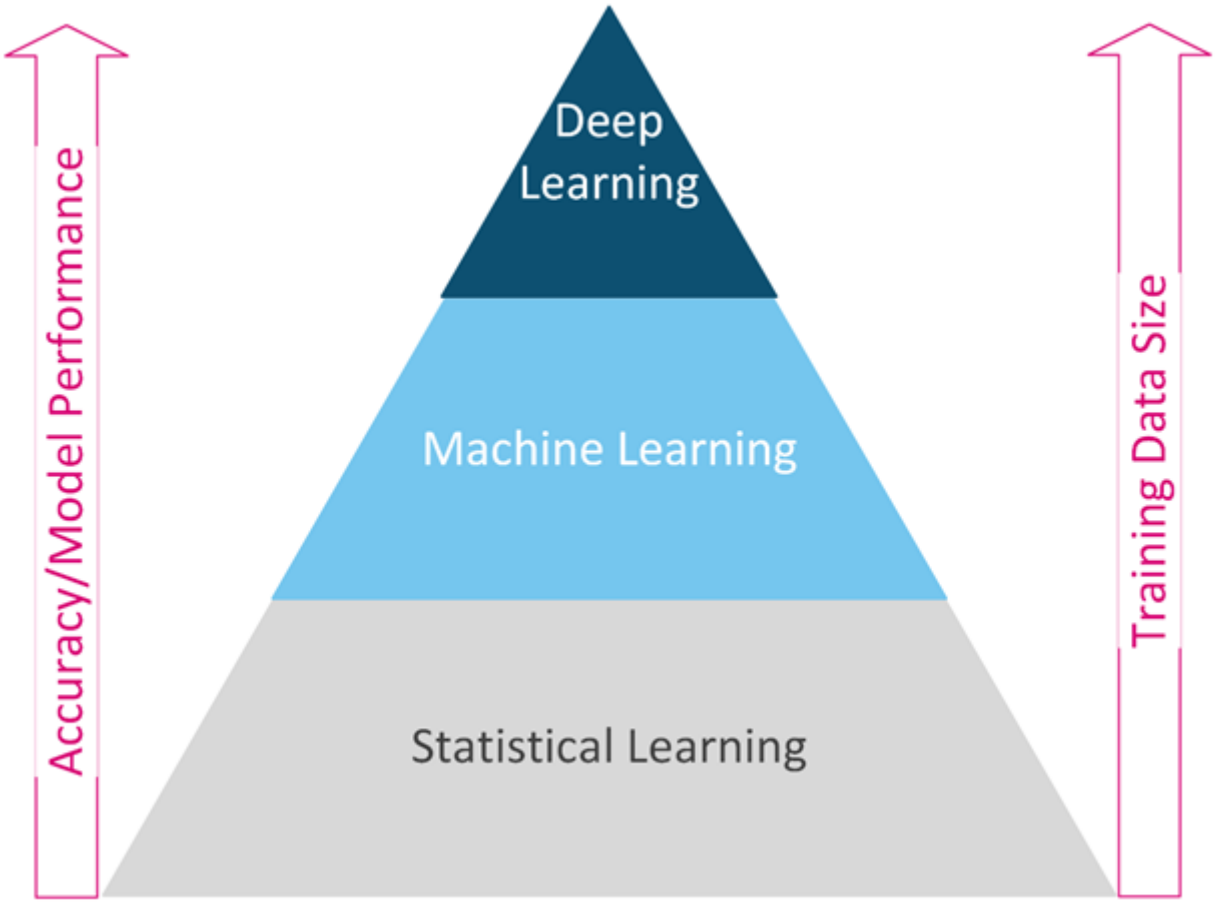

Figure 1

Model performance with data size for the three (SL, ML, DL) fields

In this section, the deep learning architectures mentioned in the article will be explained in detail.

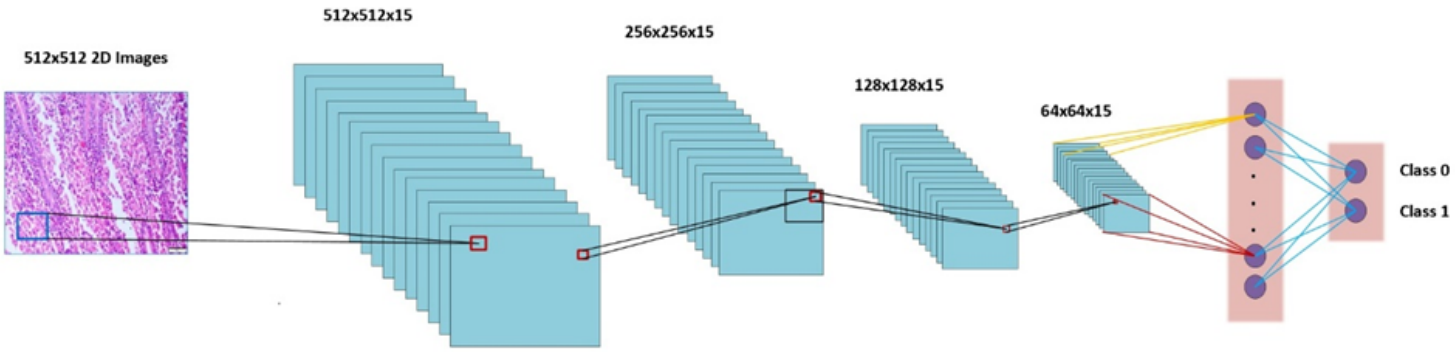

Figure 2

CNN Architecture

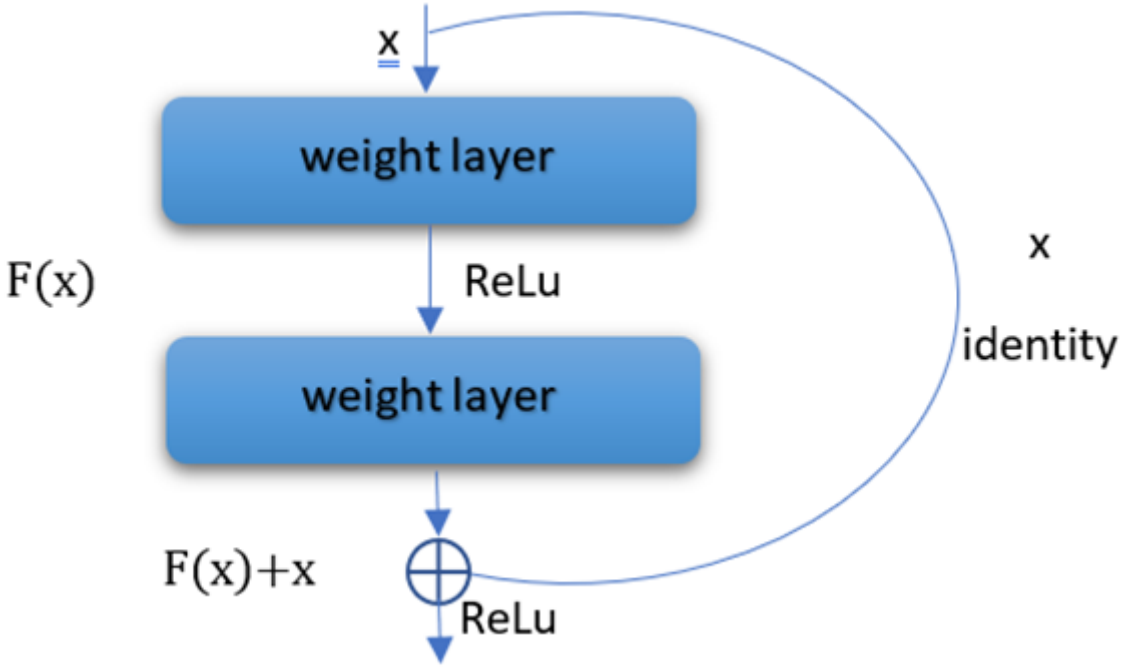

Figure 3

ResNet50 Architecture

Figure 4

The basic building block of the VGG16 network

Figure 5

a)Gut from a sheep with paratuberculosis, numerous macrophages (arrows) in submucosae, **b)** healthy intestine, H&E, scale bars=50µm.

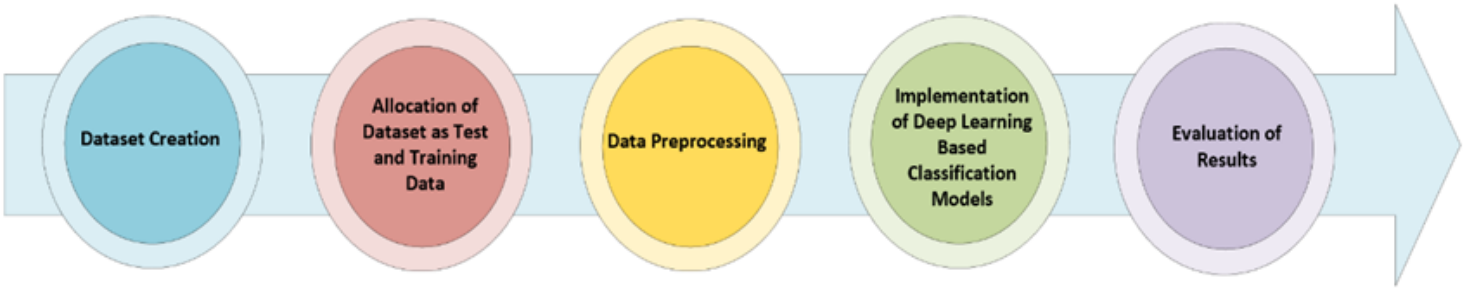

Figure 6

Flow Diagram of System

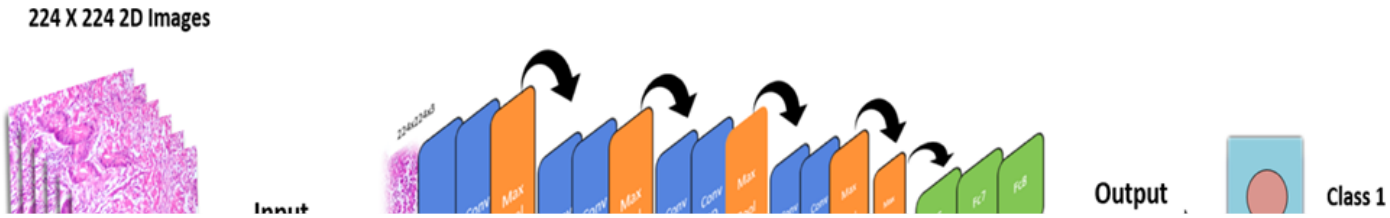

Figure 7

Proposed Fastai/VGG-16 with Grad-CAM Diagram