applied sciences

MDPI

Article

Image Visualization and Classification Using Hydatid Cyst Images with an Explainable Hybrid Model

Muhammed Yildirim

Department of Computer Engineering, Malatya Turgut Ozal University, 44200 Malatya, Türkiye; muhammed.yildirim@ozal.edu.tr

Abstract: Hydatid cysts are most commonly found in the liver, but they can also occur in other body parts such as the lungs, kidneys, bones, and brain. The growth of these cysts occurs through the division and proliferation of cells over time. Cysts usually grow slowly, and symptoms are initially absent. Symptoms often vary in size, location, and the affected organ. Common symptoms include abdominal pain, vomiting, nausea, shortness of breath, and foul odor. Early diagnosis and treatment are of great importance in this process. Therefore, computer-aided systems can be used for early diagnosis. In addition, it is very important that these cysts can be interpreted more easily by the specialist and that the error is minimized. Therefore, in this study, data visualization was performed using Grad-CAM and LIME methods for easier interpretation of hydatid cyst images via a reanalysis of data. In addition, feature extraction was performed with the MobileNetV2 architecture using the original, Grad-CAM, and LIME applied data for the grading of hydatid cyst CT images. The feature maps obtained from these three methods were combined to increase the performance of the proposed method. Then, the Kruskal method was used to reduce the size of the combined feature map. In this way, the size of the 2416×3000 feature map was reduced to 2416×700. The accuracy of the proposed model in classifying hydatid cyst images is 94%.

Keywords: artificial intelligence; CNN; deep learning; hydatid cyst; machine learning

check for updates

Citation: Yildirim, M. Image Visualization and Classification Using Hydatid Cyst Images with an Explainable Hybrid Model. *Appl. Sci.* **2023**, *13*, 9926. https://doi.org/10.3390/app13179926

Academic Editor: Andrea Prati

Received: 22 August 2023
Revised: 29 August 2023
Accepted: 31 August 2023
Published: 2 September 2023

1. Introduction

The Echinococcus granulosus worm can infect cysts, leading to a condition known as a hydatid cyst. This infection is caused by food and water contaminated with fecal material released by dogs and other carnivorous animals. This infection also occurs in other mammals [1]. A hydatid cyst is a fluid-filled sac caused by a single-celled parasite. These cysts usually show no symptoms and can grow for years without showing symptoms. However, large cysts may reduce the patient's quality of life by pressing on the surrounding tissues, or they may cause serious health problems as a result of the eruption of the cyst [2,3].

For diagnosis, X-ray, ultrasound, or magnetic resonance imaging tests are usually conducted after a physical exam. Blood tests can also be helpful for diagnosis. Treatment usually depends on the location of the cyst, the patient's general health, and the size of the cyst [4,5]. Hydatid cyst treatment may vary depending on the size of the cyst and the location of the cyst. Small cysts can be treated with medication, while large cysts usually require surgery. Medication can often be used before or after surgical treatment. The most common treatment is to remove the tumor, but sometimes medication is also recommended. Hydatid cysts can be prevented by preventing infection. It is possible to prevent the spread of infection through good hygiene habits and using enough hot water when cooking meat [6,7].

In conclusion, hydatid cysts are a common type of parasitic infection that can cause serious health problems. However, medical treatment is available and early diagnosis and treatment are important. To reduce the misclassification rates of these cysts, the image visualization and classification process is emphasized.

The classification of hydatid cyst types has been studied in some of the literature. Gul et al. carried out a classification of hydatid cyst images, initially conducting feature extraction from their dataset. The INCA (iterative neighborhood component analysis) method was used to eliminate unnecessary features from these extracted features. Then, they classified the feature map optimized using the INCA method in classical classifiers. The model the researchers used in the study had a 92% accuracy rate [8]. In this study, the dataset shared by Gul et al. was used. In the study of Gul et al., the image visualization step was not included. In this study, image visualization was performed with gradient-weighted class activation mapping (Grad-CAM) and locally interpretable model-agnostic explanation (LIME) techniques.

Wu et al. aimed to classify five different hepatic cystic echinococcosis subtypes based on deep learning algorithms using ultrasound images. Researchers used pretrained VGG19, Inception-v3, and ResNet18 models in their study. The study used a total of 1820 abdominal ultrasound pictures from 967 patients, representing five subtypes of hepatic cystic echinococcosis. In pretrained models, accuracy rates between 88.2% and 90.6% were attained. In the investigation, the researchers' highest accuracy result for the VGG19 architecture was 90.6% [9].

To categorize CT images of liver hydatid cysts, Abdulkeyim et al. employed and contrasted both conventional machine learning techniques and deep learning techniques. In the process of classification of CT images, firstly, the preprocessing step was carried out. Median filtering was used in this step. After feature extraction, decision trees and a support vector machine (SVM) were used for classification. Researchers have stated that the SVM outperforms the decision tree for CT images of hepatic hydatid disease [10].

Xin et al. stated in their study that hepatic echinococcosis is a liver disease that requires a definitive diagnosis and appropriate treatment. To reliably analyze lesions of hepatic echinococcosis, they suggested a new automated lesion segmentation and classification network that contains modules for lesion site location and lesion site segmentation. The probability map of the lesion distribution and the lesion's location was obtained by the researchers using the LRP module. CNN networks were preferred in the classification process. In the dataset consisting of CT images of 160 patients, the researchers obtained 82.45% accuracy values [11].

Image visualization is a branch of computer science that involves analyzing, processing, and displaying digital images. Image visualization technology can process pictures, video recordings, maps, and other similar digital data [12]. This process makes the data visually understandable and provides a better understanding of these images. It is used in image visualization, medicine, communications, games, cinema, and many other industries. Especially when combined with big data management and analysis, image visualization can facilitate data analysis and provide a faster understanding of images. In this study, Grad-CAM and LIME techniques were used to interpret hydatid cyst CT images more easily by experts and to prevent mistakes. In addition, these images visualized using Grad-CAM and LIME techniques were used in the developed model. The same image's various attributes were acquired in this manner. This improved the generated model's accuracy.

Artificial intelligence is used in many industries today. These areas are health, finance, education, production, energy, transportation, and defense. Health is one of the most widely used areas of artificial intelligence. In this study, artificial intelligence techniques were used to classify hydatid cyst images. It is possible to summarize the innovations and contributions of the study as follows:

- In this study, artificial intelligence techniques were used to classify hydatid cyst images. The developed model obtained more successful results than similar studies and pretrained models in the literature. The accuracy of the model developed in the classification of hydatid cyst images was 94 percent. This value indicates that the proposed model can be applied to the categorization of CT images of hydatid cysts.

- Image visualization process was carried out so that hydatid cyst images can be interpreted more easily by experts and to prevent mistakes.
- Grad-CAM and LIME techniques are used for image visualization, in which CNN architectures are used as the base.

In the second part of the article, data visualization, dataset, deep learning architectures, the Kruskal algorithm, classifiers, and the model developed for the classification of hydatid cyst images are examined. In the third chapter, the results obtained from the original dataset and the results of the proposed model are presented. In the last sections, the discussion and conclusion are provided.

2. Materials and Methods

For easier interpretation of hydatid cyst images by experts, first of all, the dataset was examined, and data visualization steps are discussed in this section. In addition, the methods used in this study and the model developed for the classification of hydatid cyst images are examined in this section.

2.1. Dataset and Image Visualization

Hydatid cyst treatment can be performed with methods such as antiparasitic drugs or surgical intervention. The dataset used in this paper was downloaded from the Kaggle platform [8,13]. This dataset consists of 5 classes. Image examples for each class in the dataset are provided in Figure 1.

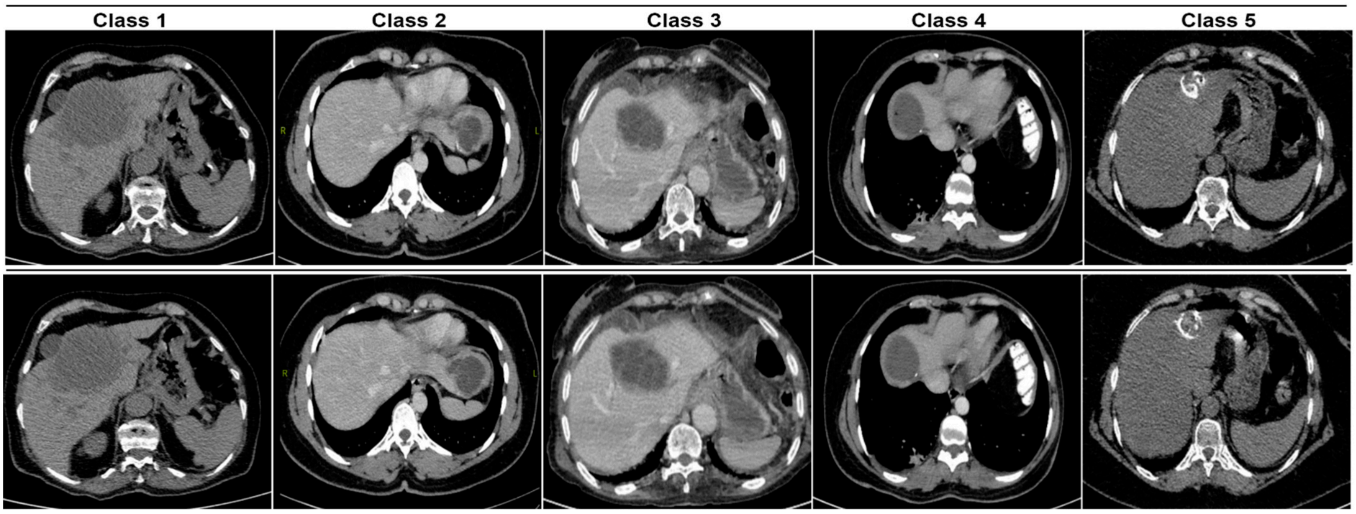

Figure 1. Hydatid cyst CT images.

There are 251 CT images in the 1st class, 541 in the 2nd class, 444 in the 3rd class, 442 in the 4th class, and 738 in the 5th class.

Image visualization is a process performed to better understand and interpret data. This process can be carried out using different methods. It is also frequently used in areas such as image visualization, signal processing, data analysis, and model building. This process allows for a better understanding of the data and a more successful decision-making process. In this study, CT images were performed using Grad-CAM and LIME technology. The images obtained after image visualization with these two techniques were used in the proposed model.

Grad-CAM is a popular research tool used to understand the importance of classification in deep learning methods. Also, thanks to this technique, it becomes possible to understand what features a classification model learns and how it makes decisions. Grad-CAM is a tool used as a visualization technique for CNNs. CNNs often make accurate predictions, but it is difficult to understand what goes on behind the model's decision-making process. With Grad-CAM technology, it becomes possible to understand which features the CNNs decide by highlighting for each class.

Because Grad-CAM is a technique for visualizing what an image classification model is classifying based on, this technique uses gradients to determine which pixels are important in making the classification decision. Grad-CAM uses gradient computation methods to generate feature maps of each class. These feature maps are then transferred onto the original image and help determine where in the image the decision was made. Since Grad-CAM is based on a linear technique, it can be implemented very quickly and easily in different neural networks. Instead of using “bounding boxes” to show where a particular object is in the classified image, Grad-CAM creates a map of the pixels associated with that object. In this way, it provides more accurate, detailed, and high-quality visualizations.

Grad-CAM can be used for many different tasks such as image classification, object detection, and segmentation. In this way, it is possible to gain a great advantage in understanding which features are influential in the decision-making process of your model. Examples of images obtained after applying Grad-CAM technology to hydatid cyst CT images are presented in Figure 2.

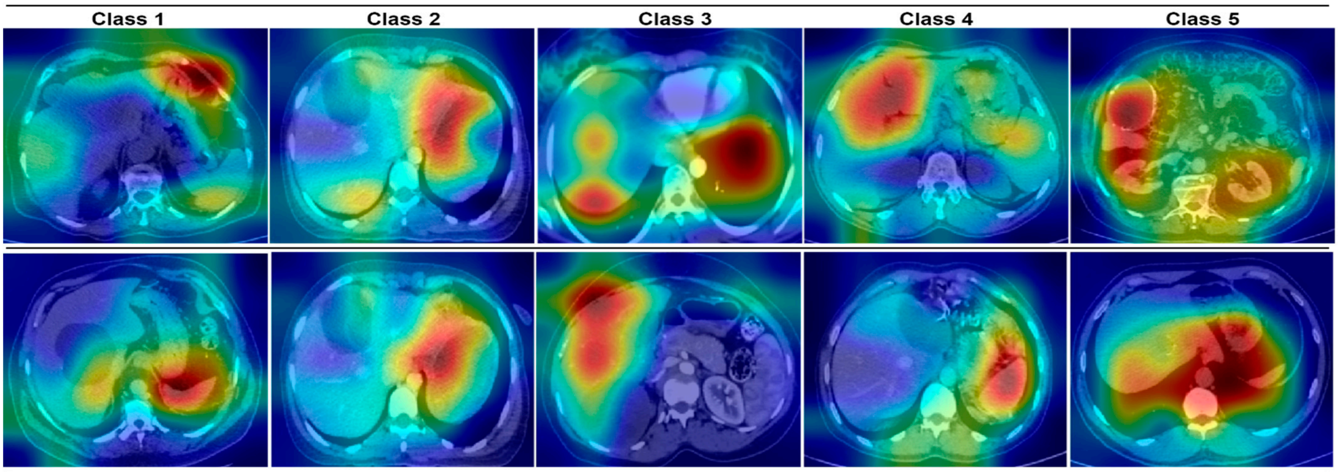

Figure 2. Image visualization with Grad-CAM technique.

Another technique used for image visualization in this study is LIME. LIME is an annotation technique used to explain the decisions of machine learning models. This technique is specifically designed for “black box” models such as complex neural network models. LIME generates random samples and compares them with individual samples to explain the model’s prediction for a given sample. It then uses the percentage of these samples to determine a measurable weight for each feature that contributes to the model’s decision. By using these weights, it becomes understandable how and why the model’s decisions are made. LIME is especially used in areas where critical decisions are made, such as medical developments, as it provides highly efficient and satisfying explanations. LIME can be used to determine which portions of an image are crucial for a particular class. In this method, classification-related accuracy issues can be improved. The maps display the portions of the image that are crucial for classification. The red areas represent the areas where the model is more concentrated. As a result, LIME creates an explanatory model using data near the input to be predicted. This explanatory model is used to clearly explain the effects of selected key features. In this way, information such as which features are affected by the outputs of a complex black box model, how they are affected, and which regions are more sensitive can be obtained. LIME is a method that makes black box models more understandable. It can be used to explain the reasons for the predictions and to make the model’s decision processes more transparent [14,15]. Examples of images obtained after applying LIME technology to hydatid cyst CT images are presented in Figure 3.

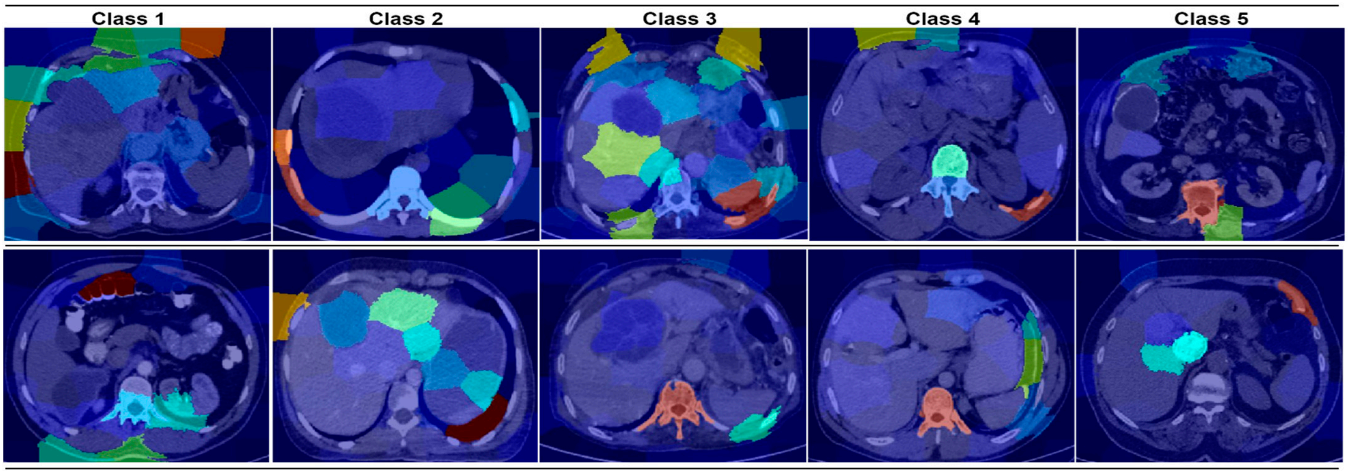

Figure 3. Image visualization with the LIME technique.

After the original hydatid cyst CT images were visualized using the Grad-CAM and LIME techniques, the obtained images were used as the basis for the proposed model.

2.2. Deep Learning, Kruskal Algorithm and Classifiers

It is possible to define deep learning as feature extraction, perception, learning, and interpretation of these features by machines. The fact that the features are not extracted by subject matter experts is one of the key characteristics that set deep learning apart from traditional machine learning. In classical machine learning methods, since the features are extracted by experts in the field, it causes time loss and increased cost. Classical machine learning methods cannot process raw data without preprocessing and expert help. With deep learning, this problem of classical machine learning has been eliminated. In other words, in deep learning architectures, features are extracted automatically thanks to the layers used. Also, there is no need to preprocess the data in deep learning architectures [16,17]. Deep learning networks consist of multiple layers that follow each other. The output of each layer is the input of the next layer or layers [18]. The rough flowchart of classical machine learning and deep learning models is presented in Figure 4.

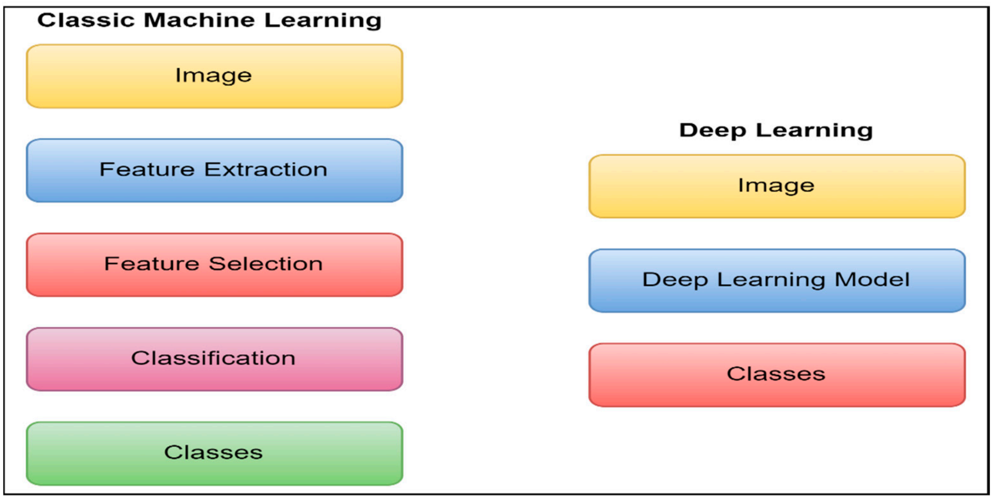

Figure 4. Flowchart of classical machine learning and deep learning models.

In this study, CNN architectures, a subbranch of deep learning, were used for automatic feature extraction. CNN architectures are one of the best-known deep learning algorithms and were developed for the computer to process data such as images, audio, and text. CNN architecture generally consists of one or more convolutional layers, pooling layers, activation functions, and classification layers (Softmax).

CNN architectures generally consist of 2 parts. In the first part, the features are extracted. In the second part, classification is made. This structure is shown in Figure 5.

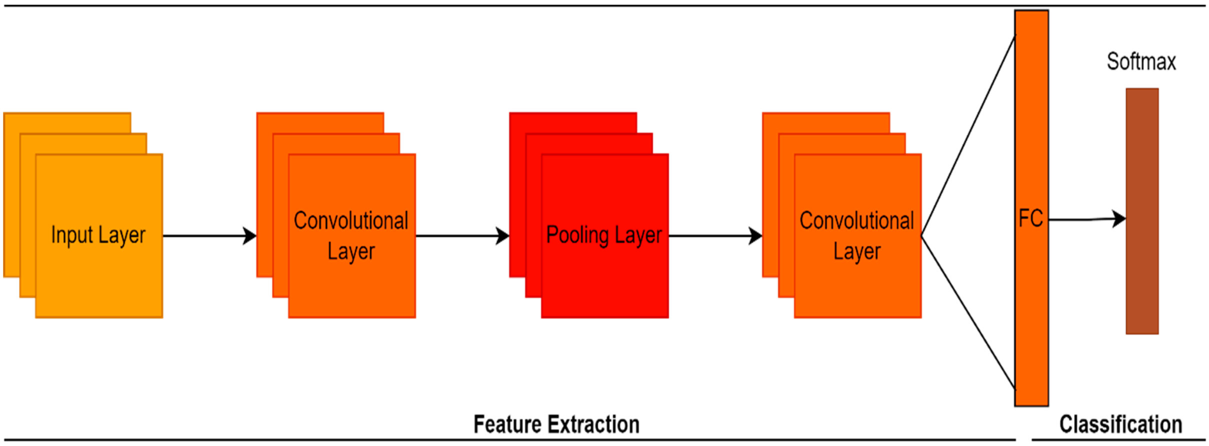

Figure 5. Feature extraction and classification process in CNN architectures.

CNN architecture provides very successful results, especially in areas such as image recognition and classification. In the proposed model, the MobileNet architecture accepted in the literature is used as the base. The MobileNet architecture was developed in 2017. It was stated by Howard et al. that although the performance rates of the developed CNN models increased, satisfactory results were not obtained in terms of speed. For this reason, researchers have tended to develop models with fewer parameters, high performance, and working speed. In this architecture, they aimed to realize mobile and deep learning applications with lower data processing capability. In this architectural feature extraction phase, instead of the classical convolution process, the deep separable convolution (depthwise separable convolutions) technique is used. This method allows for feature extraction with 8–9 times fewer parameters than the conventional convolution procedure [19].

The classification of hydatid cyst CT images was based on the MobileNetV2 architecture. Deep features were obtained from the original dataset with the MobileNetV2 architecture, from the images obtained using Grad-CAM and LIME techniques. The Kruskal algorithm was used to eliminate unnecessary features from the obtained features [20,21].

Valuable features were selected from feature maps with the Kruskal algorithm. The Kruskal dimension reduction method is a method used to find the important features in the dataset when the data are highly dimensional. In this way, fewer important features are selected while reducing the size of the dataset. These selected features are inclusive enough to describe the dataset. The Kruskal method is frequently used in data mining applications, especially classification and clustering. In addition, Kruskal is frequently used in fields such as image recognition, natural language processing, and biomedical applications. In this study, the proposed model for the classification of hydatid cyst CT images was preferred because it works faster and more effectively. Different classifiers were used to classify the feature map that was optimized using the Kruskal approach.

Machine learning classifiers are algorithms and models used to classify samples within a dataset. They use and learn from the features in the dataset to process it so they can classify future samples into the right classes. The classifiers used in the study are explained in order.

Naive Bayes: Naive Bayes classifier is a classification algorithm widely used in machine learning and statistics. This algorithm is based on Bayes' theorem and uses learning data to classify. The naive Bayes classifier calculates the relationship of the properties of the observations (inputs) with the classes according to a certain probability distribution. This probability distribution is used to calculate the probability that the given inputs belong to each class, and as a result, the class with the highest probability is selected [21].

Decision trees: Decision trees are a graphical representation that helps classify a dataset based on a set of criteria. These criteria are based on the characteristics of the data and classification is made by creating a decision point for each feature level. The decision trees algorithm is used to classify datasets and all results are displayed as a tree structure. First, a tree structure is created according to the characteristics of the data, and then the dataset is classified using this tree structure [22].

K-nearest neighbors (KNN): KNN classifies new data points based on the classification of the K closest samples in the dataset in which it is located. It is generally used in technologies such as bioinformatics and fingerprint recognition. The KNN classifier is widely used, especially in the fields of big data and data mining. The KNN classifier does not require the data to be linearly separable. Also, KNN does not require any predictive model or parameter selection during the training, testing, or validation phase [23].

Logistic regression: Logistic regression is a classification algorithm. The output is given the probability that a data point belongs to one of the two classes. Usually, the classification process is performed using a linear regression model. However, linear regression is not a suitable method for classification problems because the output value is a continuous variable and cannot predict values outside the bounds. Logistic regression uses a linear regression model for classification while giving a class label in the final output [24].

Support vector machine (SVM): SVM is a learning algorithm used to classify samples in a dataset into different classes. SVM divides the samples in the dataset into 2 or more classes and selects the best-separated plane or hyperplane when classifying. SVM is a classification method that works efficiently, especially in high-dimensional feature spaces. This is because SVM uses a kernel function that moves the feature space to a low-dimensional subspace and then chooses a bounding plane to differentiate between classes [25].

Ensemble subspace KNN (ES): This method first divides the data into subfeature spaces and then creates a classification model in each subfeature space using the KNN method. The results of these submodels are then aligned and combined. The purpose of this method is to increase the classification performance by using more than one sub-model in cases where the data may behave differently in structurally different feature spaces. This technique is particularly useful as it is more scalable for higher dimensional data due to shrinking the feature space and reducing the overfitting problem [26,27].

2.3. Proposed Hydatid Cyst Model

A new model was developed in this study for the classification of hydatid cyst CT images. In the developed model, MobileNetV2 architecture is used as the base [28]. While feature extraction is carried out by an expert in hand-crafted methods, feature extraction is performed automatically in CNN architectures. Therefore, feature extraction in the proposed model is taken from the logits layer of the MobileNetV2. In order to increase the performance metrics of the proposed model, the data visualization step was also used. Image visualization was performed using Grad-CAM and LIME. Using MobileNetV2 architecture, 2416×1000 feature original, 2416×1000 feature Grad-CAM-applied images, and 2416×1000 feature LIME-applied images were obtained. In this step, different features of the same image are highlighted. The feature extraction stage is presented in Figure 6.

The feature maps obtained after feature extraction were then combined. After the merge, the size of the new feature map was 2416×3000. In order for the proposed model to run faster, the Kruskal algorithm was applied from the feature map of 2416×3000 size. The size of the feature map optimized with the Kruskal algorithm was 2416×700. After this optimization step, the size of the feature map was reduced. In this step, the aim is that the model runs faster and achieves more successful results. In the last step of the proposed model, the optimized feature map of 2416×700 size was classified using 6 different classifiers. The feature aggregation, size reduction, and classification steps are presented in Figure 7.

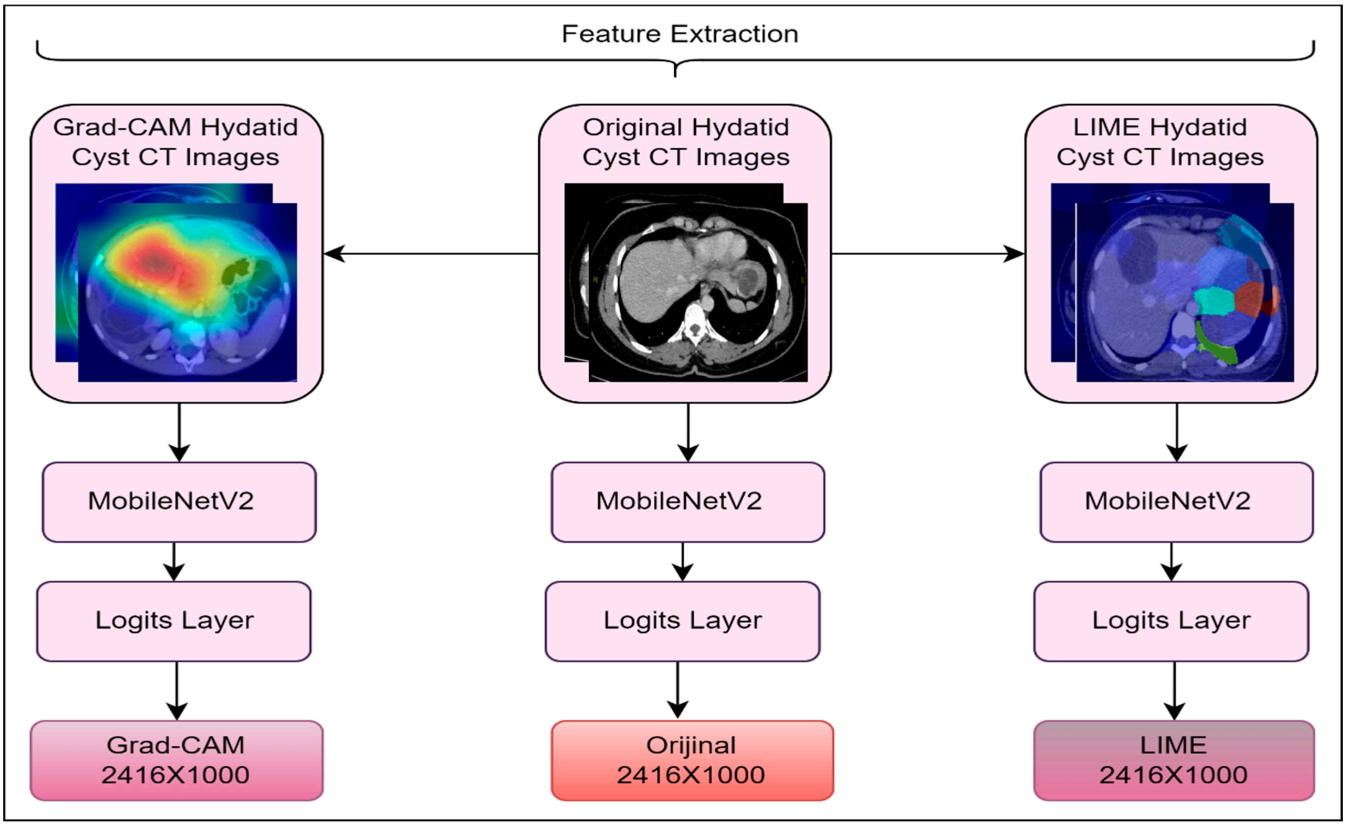

Figure 6. Feature extraction.

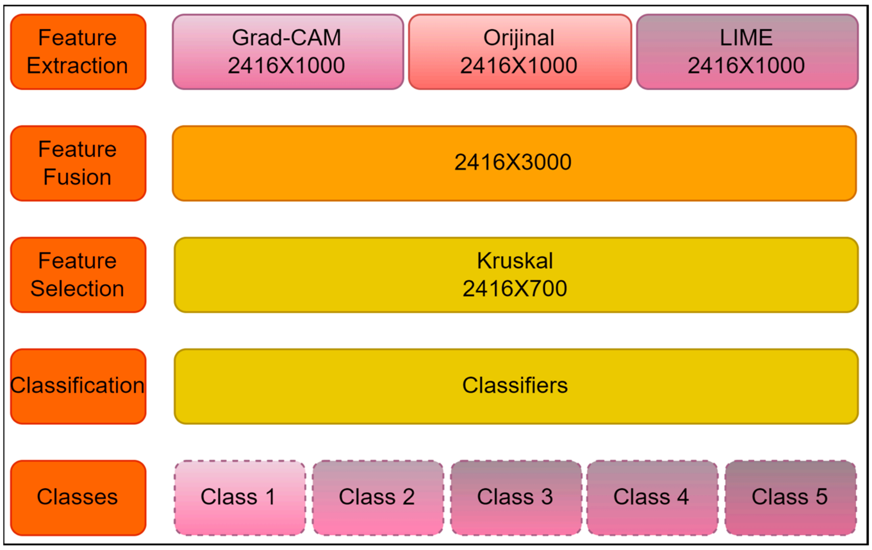

Figure 7. Flowchart of the proposed hydatid cyst model.

In the final stage of the proposed hydatid cyst model, the relevant image is included in any of the 5 classes since there are 5 classes in this study.

3. Results

In this study, which was carried out for the classification of hydatid cyst CT images, the application results were taken in a MATLAB environment. In all models, the results were obtained under the same conditions, and 80 percent of the dataset was used for training and 20 percent for testing. In order to compare the performance of the proposed model, separate results were obtained from the original dataset, and from the Grad-CAM- and

LIME-applied datasets. While obtaining the results from the original dataset, the feature map was extracted using the MobileNetV2 architecture, the extracted feature map was then optimized with the Kruskal method, and the optimized feature map was classified in six different classifiers. The flowchart of the proposed model is presented in Figure 8 while obtaining results from the original images.

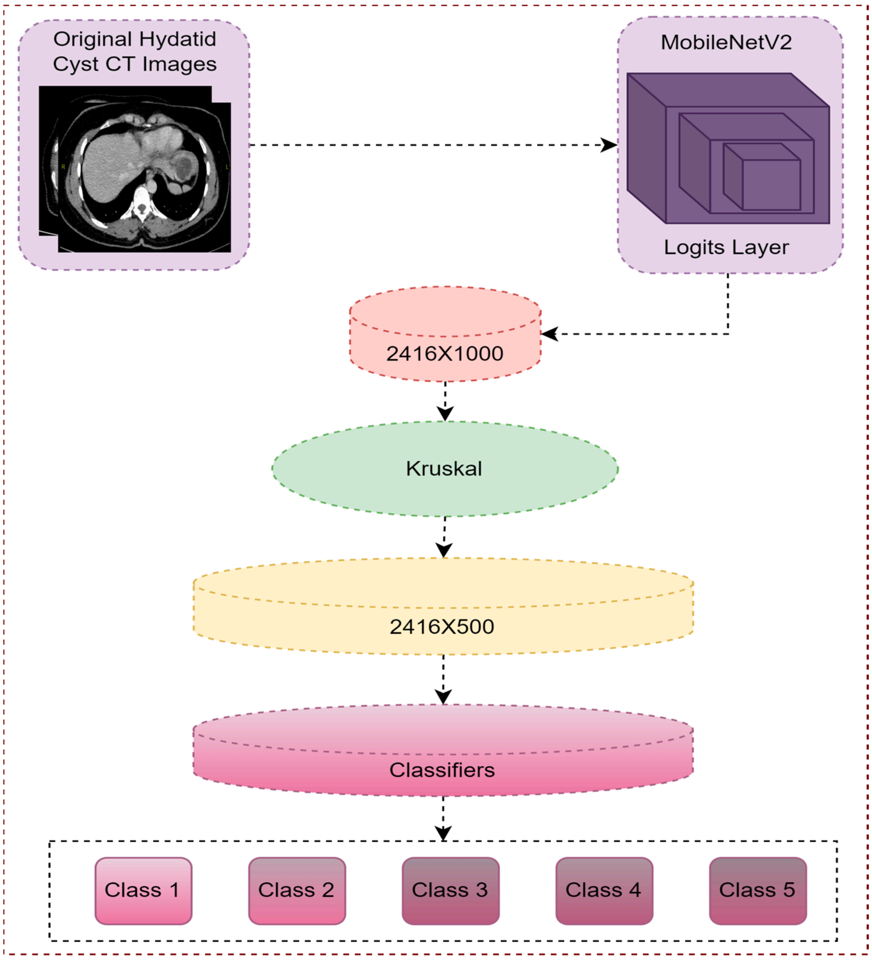

Figure 8. Flowchart of the proposed model in the original dataset.

As seen in Figure 8, MobileNetV2 architecture was used for feature extraction from the images in the original dataset. Feature maps were performed using the logits layer of the MobileNetV2 architecture. While the size of the feature map taken from this layer was 2416 $\times$ 1000, after the Kruskal process, the new size became 2416 $\times$ 500. The Kruskal method's feature map was finally classified using six different classifiers. Confusion matrices obtained in these classifiers are presented in Figure 9.

When Figure 9 is reviewed, it can be observed that the KNN classifier is the most effective classifier, and the FT classifier is the least effective classifier. Of the 483 hydatid cyst CT images, FT predicted 243 correctly and 240 incorrectly. The KNN, which achieved the highest accuracy in classifying hydatid cyst CT images, predicted 448 correctly and 35 incorrectly predicted 483 hydatid cyst CT images. KNN predicted 47 correctly and 3 incorrectly out of 50 test images belonging to Class 1. It predicted 103 of 108 test images belonging to Class 2 correctly and 5 of them incorrectly. KNN classifiers predicted 84 correctly and 5 incorrectly out of 89 test images belonging to Class 3. It predicted 77 correctly and 12 incorrectly out of 89 test images belonging to Class 4. It predicted 137 correctly and 10 incorrectly out of 147 test images belonging to Class 5. The feature map obtained from the logits layer of the MobileNetV2 architecture was optimized with the Kruskal algorithm; the accuracy values obtained in six different classifiers are presented in Table 1.

FT (True Class)	1	2	3	4	5
1	8	15	2	11	14
2	10	60	10	15	13
3	7	14	46	17	5
4	4	11	6	37	31
5	11	17	5	22	92

Predicted Class

NB (True Class)	1	2	3	4	5
1	28	8	7	3	4
2	6	76	12	8	6
3	6	6	57	10	10
4	9	11	1	55	13
5	9	10	11	27	90

Predicted Class

LD (True Class)	1	2	3	4	5
1	41	5	0	3	1
2	4	86	7	5	6
3	1	3	79	1	5
4	3	6	1	68	11
5	2	7	4	10	124

Predicted Class

SVM (True Class)	1	2	3	4	5
1	39	6	0	3	2
2	2	99	4	1	2
3	0	6	79	1	3
4	1	3	1	77	7
5	0	0	3	3	141

Predicted Class

ES (True Class)	1	2	3	4	5
1	46	2	0	2	0
2	1	103	3	1	0
3	0	4	84	0	1
4	1	2	2	77	7
5	0	2	1	7	137

Predicted Class

KNN (True Class)	1	2	3	4	5
1	47	1	0	2	0
2	1	103	3	1	0
3	0	4	84	0	1
4	1	2	2	77	7
5	0	3	2	5	137

Predicted Class

Figure 9. Confusion matrices obtained in the original dataset.

Table 1. Accuracy values reached in the original dataset (%).

FT	NB	LD	SVM	ES	KNN
50.3	63.4	84.2	90.1	92.5	92.8

When Table 1 is reviewed, it can be noted that the KNN classifier achieved the greatest accuracy value at 92.8%, and the FT classifier obtained the lowest accuracy value at 50.3%.

The proposed hybrid model for the classification of hydatid cyst CT images was first extracted from original, LIME, and Grad-CAM images using the MobileNetV2 architecture, and 1000 features were extracted for each image. The size of the feature map obtained after combining the extracted features was 2416 × 3000. The new dimension of the feature map optimized using the Kruskal method is 2416 × 700. The suggested model functions more quickly and efficiently in this fashion. Finally, the optimized feature map was classified in six different classifiers. Confusion matrices obtained in these classifiers are presented in Figure 10.

KNN (True Class)	1	2	3	4	5
1	47	2	1	0	0
2	1	103	0	3	1
3	0	2	85	0	2
4	3	1	2	82	0
5	0	0	1	8	139

Predicted Class

ES (True Class)	1	2	3	4	5
1	48	0	0	2	0
2	3	103	1	2	0
3	1	1	83	1	2
4	0	0	1	81	6
5	1	2	1	5	139

Predicted Class

SVM (True Class)	1	2	3	4	5
1	39	4	0	3	4
2	1	101	2	1	4
3	1	9	77	1	0
4	0	2	0	80	6
5	2	1	1	10	134

Predicted Class

LD (True Class)	1	2	3	4	5
1	35	10	0	2	3
2	8	95	3	1	2
3	0	4	78	2	4
4	0	2	4	69	13
5	4	6	0	19	119

Predicted Class

NB (True Class)	1	2	3	4	5
1	30	8	0	5	7
2	5	85	6	8	5
3	9	14	51	10	4
4	2	8	1	63	14
5	8	9	5	36	90

Predicted Class

FT (True Class)	1	2	3	4	5
1	11	12	5	8	14
2	5	57	7	23	17
3	5	9	41	17	16
4	5	9	12	40	22
5	11	7	14	28	88

Predicted Class

Figure 10. Confusion matrices obtained in the proposed hydatid cyst model.

When Figure 10 is inspected, it can be observed that the KNN classifier is the most effective classifier, and the FT classifier is the least effective classifier. Of the 483 hydatid cyst CT images, the FT predicted 237 correctly and 242 incorrectly. KNN, which achieved the highest accuracy in classifying hydatid cyst CT images, predicted 456 of 483 hydatid cyst CT images correctly and 27 of them incorrectly. KNN predicted 47 correctly and 3 incorrectly out of 50 test images belonging to Class 1. It predicted 103 of 108 test images belonging to Class 2 correctly and 5 of them incorrectly. KNN predicted 85 correctly and 4 incorrectly out of 89 test images belonging to Class 3. KNN predicted 82 correctly and 6 incorrectly out of 88 test images belonging to Class 4. KNN predicted 139 correctly and 9 incorrectly out of 148 test images belonging to Class 5. The KNN classifier was the most successful among the classifiers used in this study to classify hydatid cyst images because the KNN method is a simple and understandable classification algorithm. Compared to the training and implementation processes of other complex machine learning algorithms, KNN is easier to implement and understand. KNN relies on data without making any biases or assumptions. This allows KNN to offer a flexible classification approach, given that hydatid cyst images can be complex and variable in nature. The accuracy values reached by the proposed model in six different classifiers are shown in Table 2.

Table 2. The classifier accuracy values of the proposed model (%).

FT	NB	LD	SVM	ES	KNN
49.1	66	82	89.2	94	94.4

In the proposed model, the fine KNN classifier's greatest accuracy value was attained. The performance metrics obtained in the fine KNN classifier and the training times of the models are presented in Table 3.

The proposed model achieves the maximum accuracy in Class 5 when Table 3 is reviewed. Class 5 images classification accuracy of the proposed model was 97.88%. Class 3 was the group in which the suggested model failed the most. Class 3 image classification accuracy of the proposed model was 93.18%.

The AUC/ROC curve obtained in the proposed hydatid cyst model is in Figure 11.

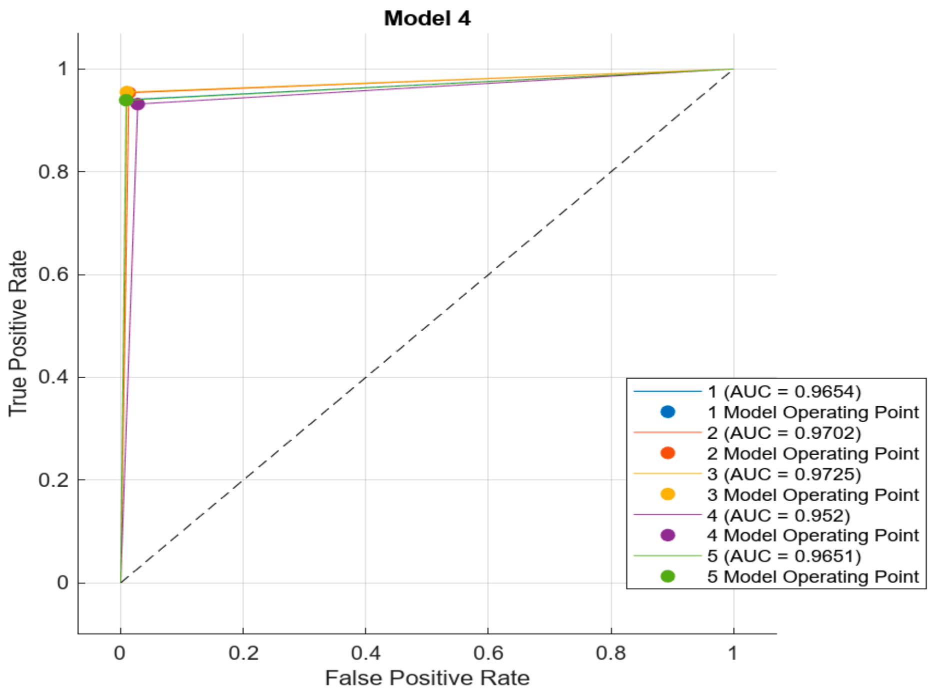

Figure 11. AUC/ROC curves of proposed hydatid cyst model.

Table 3. Performance values of the proposed hydatid cyst model (%).

	Accuracy	Sensitivity	Specificity	FPR	FDR	FNR	F1
Class 1	94	92.15	99.30	0.69	6	7.84	93.06
Class 2	95.37	95.37	98.66	1.33	4.62	4.62	95.37
Class 3	95.50	95.50	98.98	1.01	4.49	4.49	95.50
Class 4	93.18	88.17	98.46	1.53	6.81	11.82	90.60
Class 5	97.88	93.91	99.10	0.89	2.11	6.08	95.86

As can be seen in Figure 11, it has been observed that the proposed hydatid cyst model has achieved a high performance.

4. Discussion

Hydatid cyst disease is an infection caused by the Echinococcus granulosus parasite, which is endemic in many countries. The disease causes cyst formation in many organs of the body. It is a common disease, especially in small cattle-breeding countries, including our country. It is transmitted to humans by ingesting egg-contaminated food from hosts such as dogs, foxes, and sheep.

In adults, the cyst most often settles in the liver. Imaging methods such as ultrasonography (USG), computed tomography (CT), and magnetic resonance (MRI) are used for diagnosis [29,30]. CT shows the presence of calcification in the cyst. It is also used in the evaluation of complications. Intrahepatic and extrahepatic complications of hydatid cyst disease can be seen. Intrahepatic complications are infection, biliary tract opening, and vascular findings. Extrahepatic complications are hematogenous spread to distant organs and peritoneal rupture [31,32].

Hydatid cyst is a disease with high morbidity with cyst formation and complications in many organs in the body. Radiological imaging has an important role in the diagnosis, treatment, and follow-up of this disease. With imaging methods, it is determined whether the cyst is active or inactive, and the treatment steps are decided.

Hydatid cyst infection is a serious health problem, and early diagnosis and treatment are important. Early detection can reduce the growth of cysts and damage to surrounding tissues and reduce the risk of further infections. It can also help prevent serious complications that can have fatal consequences if left untreated. Symptoms of hydatid cyst infection are usually very mild and may go unnoticed in the early stages of the disease [33,34]. Therefore, in this study, first of all, the image visualization process was carried out so that the hydatid cyst CT images can be interpreted more easily and error-free by experts. Then, hydatid cyst CT images were classified. The model developed in the study obtained better results than the pretrained models accepted in the literature and similar studies in the literature.

There are similar studies in the literature for the classification of hydatid cyst CT images. Studies on the classification of hydatid cyst CT images are presented in Table 4.

Table 4. Literature review on hydatid cyst CT images.

Study	Year	Method	Number of Patients	Accuracy
Wu et al. [9]	2022	VGG19, InceptionV3, ResNet18	967	90.6%
Abdulkeyim et al. [10]	2019	Conventional machine learning, deep learning	-	-
Xin et al. [11]	2020	CNN	160	82.45%
Gul et al. [8]	2023	CNN, INCA	2416	92.0%
Proposed model	2023	MobileNetV2, Grad-CAM, LIME, classifiers	2416	94.0%

When similar studies in the literature for the classification of hydatid cyst CT images are examined, it is seen that the proposed model is successful. The accuracy rates we obtained in this study are for the classification of liver hydatid disease, and additional studies are needed for its application for cysts in different localizations in the body.

The biggest limitation of this study is the collection of data from a single center. In future studies, it is among our goals to collect data from more centers and design systems that work online.

5. Conclusions

The need to reduce costs, save time, and improve efficiency is apparent given the heavy patient load and workload in hospitals today. Due to all of these factors, using artificial intelligence techniques to evaluate medical images offers a significant benefit. The suggested model is anticipated to play a significant role in the categorization of CT images of hydatid cysts. The hybrid model suggested in this paper produced an accuracy rate of 94%. This result demonstrates the applicability of the suggested hybrid technique for categorizing CT images of hydatid cysts.

Funding: This research received no external funding.

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Soltany, S.; Hemmati, H.R.; Toussy, J.A.; Nazifi, H.; Alibakhshi, A.; Toosi, P.A. Evaluation of musculoskeletal hydatid cyst cases in terms of clinical manifestations, method of dealing, treatment, and recurrence. *Open Access Maced. J. Med Sci.* **2020**, *8*, 99–104. [CrossRef]
2. Chávez-Ruvalcaba, F.; Chávez-Ruvalcaba, M.I.; Santibañez, K.M.; Muñoz-Carrillo, J.L.; Coria, A.L.; Martínez, U.R. Foodborne Parasitic Diseases in the Neotropics—A review. *Helminthologia* **2021**, *58*, 119–133. [CrossRef]
3. Khemasuwan, D.; Farver, C.F.; Mehta, A.C. Parasites of the Air Passages. *Chest* **2014**, *145*, 883–895. [CrossRef]
4. Acar, A.; Rodop, O.; Yenilmez, E.; Baylan, O.; Oncül, O. Case report: Primary localization of a hydatid cyst in the adductor brevis muscle. *Turk. J. Parasitol.* **2009**, *33*, 174–176.
5. Padayachy, L.; Ozek, M. Hydatid disease of the brain and spine. *Child's Nerv. Syst.* **2022**, *39*, 751–758. [CrossRef] [PubMed]
6. Bayrak, M.; Altıntas, Y. Current approaches in the surgical treatment of liver hydatid disease: Single center experience. *BMC Surg.* **2019**, *19*, 95. [CrossRef] [PubMed]
7. Das, A.K.; Nanda, P.K.; Das, A.; Biswas, S. Hazards and safety issues of meat and meat products. In *Food Safety and Human Health*; Academic Press: Cambridge, MA, USA, 2019; pp. 145–168.
8. Gul, Y.; Muezzinoglu, T.; Kilicarslan, G.; Dogan, S.; Tuncer, T. Application of the deep transfer learning framework for hydatid cyst classification using CT images. *Soft Comput.* **2023**, *27*, 7179–7189. [CrossRef]
9. Wu, M.; Yan, C.; Wang, X.; Liu, Q.; Liu, Z.; Song, T. Automatic classification of hepatic cystic echinococcosis using ultrasound images and deep learning. *J. Ultrasound Med.* **2022**, *41*, 163–174. [CrossRef]
10. Abdukeyim, N.; Liu, Z.; Yan, C. CT Images Classification of Hepatic Hydatid Disease by Different Methods. In Proceedings of the 2019 6th International Conference on Systems and Informatics (ICSAI), Shanghai, China, 2–4 November 2019; IEEE: New York, NY, USA, 2019.
11. Xin, S.; Shi, H.; Jide, A.; Zhu, M.; Ma, C.; Liao, H. Automatic lesion segmentation and classification of hepatic echinococcosis using a multiscale-feature convolutional neural network. *Med. Biol. Eng. Comput.* **2020**, *58*, 659–668. [CrossRef]
12. Zhang, Y.-J. Image engineering. In *Handbook of Image Engineering*; Springer: Singapore, 2021; pp. 55–83.
13. Available online: https://www.kaggle.com/datasets/tahamu/hydatid-cyst?select=2 (accessed on 20 June 2023).
14. Ribeiro, M.T.; Singh, S.; Guestrin, C. Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.
15. Tasci, B.; Tasci, I. Deep feature extraction based brain image classification model using preprocessed images: PDRNet. *Biomed. Signal Process. Control.* **2022**, *78*, 103948. [CrossRef]
16. Jiménez-Luna, J.; Grisoni, F.; Weskamp, N.; Schneider, G. Artificial intelligence in drug discovery: Recent advances and future perspectives. *Expert Opin. Drug Discov.* **2021**, *16*, 949–959. [CrossRef] [PubMed]
17. Park, T.J.; Kanda, N.; Dimitriadis, D.; Han, K.J.; Watanabe, S.; Narayanan, S. A review of speaker diarization: Recent advances with deep learning. *Comput. Speech Lang.* **2022**, *72*, 101317.

18. Attique Khan, M.; Sharif, M.; Akram, T.; Kadry, S.; Hsu, C.H. A two-stream deep neural network-based intelligent system for complex skin cancer types classification. *Int. J. Intell. Syst.* **2022**, *37*, 10621–10649. [CrossRef]
19. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
20. Malji, P.; Sakhare, S. Significance of entropy correlation coefficient over symmetric uncertainty on FAST clustering feature selection algorithm. In Proceedings of the 2017 11th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, India, 5–6 January 2017; IEEE: New York, NY, USA, 2017.
21. Vora, S.; Yang, H. A comprehensive study of eleven feature selection algorithms and their impact on text classification. In Proceedings of the 2017 Computing Conference, London, UK, 18–20 July 2017; IEEE: New York, NY, USA, 2017.
22. Quinlan, J.R. Decision trees and decision-making. *IEEE Trans. Syst. Man Cybern.* **1990**, *20*, 339–346. [CrossRef]
23. Dudani, S.A. The Distance-Weighted k-Nearest-Neighbor Rule. *IEEE Trans. Syst. Man, Cybern.* **1976**, *SMC-6*, 325–327. [CrossRef]
24. Pregibon, D. Logistic regression diagnostics. *Ann. Stat.* **1981**, *9*, 705–724. [CrossRef]
25. Burges, C.J. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]
26. Ho, T.K. Nearest neighbors in random subspaces. In *Advances in Pattern Recognition, Proceedings of the Joint IAPR International Workshops SSPR'98 and SPR'98, Sydney, Australia, 11–13 August 1998*; Springer: Berlin/Heidelberg, Germany, 1998.
27. Laaksonen, J. *Subspace Classifiers in Recognition of Handwritten Digits*; Helsinki University of Technology: Helsinki, Finland, 1997.
28. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
29. Inaebnit, E.F.; Romero, F.X.M.; Sampedro, J.J.S.; Argenté, X.G.; Canis, J.M.M. A review of the diagnosis and management of liver hydatid cyst. *Rev. Española Enfermedades Dig.* **2021**, *114*, 35–41.
30. Alshoabi, S.A.; Alkalady, A.H.; Almas, K.M.; Magram, A.O.; Algaberi, A.K.; Alareqi, A.A.; Hamid, A.M.; Alhazmi, F.H.; Qurashi, A.A.; Abdulaal, O.M.; et al. Hydatid Disease: A Radiological Pictorial Review of a Great Neoplasms Mimicker. *Diagnostics* **2023**, *13*, 1127.
31. Toumi, O.; Ammar, H.; Gupta, R.; Ben Jabra, S.; Hamida, B.; Noomen, F.; Zouari, K.; Golli, M. Management of liver hydatid cyst with cystobiliary communication and acute cholangitis: A 27-year experience. *Eur. J. Trauma Emerg. Surg.* **2019**, *45*, 1115–1119. [CrossRef] [PubMed]
32. Borahma, M.; Jebari, Y.; Chabib, F.Z.; Lagdali, N.; Benelbarhdadi, I.; Ajana, F.Z. Endoscopic treatment of a liver hydatid cyst with biliary rupture. *Endoscopy* **2021**, *54*, E456–E457. [CrossRef] [PubMed]
33. Padmane, E.; Gujar, S. Case Report on Hydatid Cysts with Hydropnemonothorax. *J. Pharm. Res. Int.* **2021**, *33*, 1–5. [CrossRef]
34. Garoma, A. Epidemiology of Hydatidosis in Ethiopia. *J. Nurs. Health Care* **2022**, *7*, 1–6.