



RESEARCH ARTICLE / ARAŞTIRMA MAKALESİ

# A Novel Brain Tumor Detection Approach: Integrating Deep Learning and Support Vector Machines

## Yeni Bir Beyin Tümörü Tespiti Yaklaşımı: Derin Öğrenme ve Destek Vektör Makinelerinin Entegre Edilmesi

Ezgi Özer 

Piri Reis University, Faculty of Engineering, Computer Engineering, İstanbul, TÜRKİYE  
 Corresponding Author / Sorumlu Yazar : ezgiozherhs@gmail.com

### Abstract

Brain tumors are among the most common causes of death in humans. Early and accurate detection of brain cancers is critical for effective treatment. Imaging techniques such as computed tomography, magnetic resonance imaging, X-rays, and ultrasound are used as a preliminary reference by illness experts. Different learning strategies have been employed in the field of health to diagnose diseases early, reduce the intensity of experts, and minimize diagnostic errors. Image processing studies in brain research have begun to provide successful findings in recent years, thanks to the developed of machine learning and deep learning models. In this study, as a novelty to the studies in the literature, a hybrid algorithm is proposed that features were extracted with pre-trained based CNN, classification was made with SVM based different kernels. As a result, the brain tumors were detected with 98% classification performance.

**Keywords:** Tumor Detection, Feature Extraction, Pre-Trained Models, Support Vector Machine

### Öz

Beyin tümörleri insanlarda en sık görülen ölüm nedenleri arasındadır. Beyin kanserlerinin erken ve doğru tespiti etkili tedavi için kritik öneme sahiptir. Bilgisayarlı tomografi, manyetik rezonans görüntüleme, X-ışınları ve ultrason gibi görüntüleme teknikleri hastalık uzmanları tarafından ön referans olarak kullanılmaktadır. Hastalıkları erken teşhis etmek, uzmanların yoğunluğunu azaltmak ve teşhis hatalarını en aza indirmek için sağlık alanında farklı öğrenme stratejileri kullanılmaktadır. Beyin araştırmalarında görüntü işleme çalışmaları son yıllarda makine öğrenmesi ve derin öğrenme modellerinin geliştirilmesi sayesinde başarılı bulgular sağlamaya başlamıştır. Bu çalışmada literatürdeki çalışmalara bir yenilik olarak, önceden eğitilmiş CNN tabanlı özellikler çıkarımı yapılan, SVM tabanlı farklı çekirdeklerle sınıflandırma yapılan bir hibrit algoritma önerilmiştir. Sonuç olarak beyin tümörleri %98 sınıflandırma performansı ile tespit edilmiştir.

**Anahtar Kelimeler:** Tümör Tespiti, Özellik Çıkarımı, Önceden Eğitilmiş Model, Destek Vektör Makineleri

### 1. Introduction

Cancer is a disease in which cells grow uncontrollably and form a tissue mass. In some cases, the cells resulting from this mass spread to different parts of the body. Regular cell division, cell differentiation, and genetic changes seen in healthy tissues are among the most important causes of cancer. As a result of these changes, the factors that cause cell groups to multiply and grow increase, and at the same time, the factors that enable the destruction of diseased cells by natural means and the immune system decrease. 'Malignant' metastatic tumors, as the term frequently used to describe cancers, cause widespread spread to the region where they are located and even to distant tissues and organs, unlike benign tumors defined as benign [1]. The cells that make up cancerous tissues show largely heterogeneous properties. Although the cause of cancer is not known exactly, there are two different risk group factors that trigger cancer. These are changeable and non-changeable factors. Factors that cannot be changed are age, gender and family history, etc. while they are changeable environmental factors such as smoking and alcohol use, poor eating habits, air pollution, and radiation exposure [2]. According to the latest report published by the International Agency for Research on Cancer (IARC) by the World Health Organization (WHO), an estimated 20 million new cancer

cases were seen in 2022, and 9.7 million deaths occurred due to cancer [3].

Today, modern imaging techniques, including Magnetic Resonance (MR), X-ray Imaging, Ultrasonography, Computed Tomography, etc.), help detect cancer sensitively and quickly as a benign tumor or malignant tumor, with a deep learning-based decision support system [4]. Considering the data size, it is not enough to obtain the tumor with high success. It is also very important to make faster predictions by using fewer variables, that is, with more accurate models. This study aims to present a new artificial intelligence-based decision support system for cancer detection by examining the studies proposed in the existing literature on brain tumors. The proposed algorithm can provide high sensitivity, specificity, and accuracy, can also have the potential to contribute to early diagnosis and effective treatment in clinical practice.

### 2. Motivation and Overview

Early detection and accurate diagnosis are critical to increase the effectiveness of treatment. Imaging tools play an important role in detecting brain tumors in the early stages. Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) show abnormalities in the brain more clearly, accelerating the diagnosis and treatment processes [4]. The images obtained are

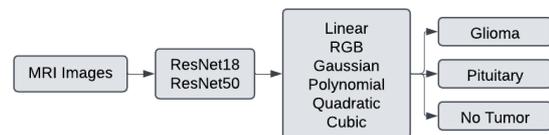
interpreted by doctors or experts. These methods can sometimes be invasive, costly, and have a risk of human error in interpreting the results. Detection of brain tumors plays a critical role in determining the course of the disease and treatment strategies. An early diagnosis system to be developed can improve the patient's quality of life and increase treatment success rates.

In recent years, different algorithms have been proposed for processing medical images and diagnosing brain tumors. Vani et al. [5] proposed a machine learning-based SVM method to classify brain tumors. They predicted positive and negative brain tumors with 82% and 81.48% accuracy, respectively. Mohsen et al. [6] proposed a new method to classify brain tumors using deep learning methods and Discrete Wavelet Transform (DWT). They achieved 93.94% accuracy with this model. Shahzadi et al. [7] used Convolutional Neural Networks (CNN)-based hybrid model structure to detect brain tumors. Feature extraction and classification used Long Short Term Memory structure along with AlexNet and VggNet CNN models. In the study, they achieved 71% accuracy with AlexNet-LSTM and 84% accuracy with VGGNet-LSTM. Swati et al. [8] proposed transfer learning for classification of brain tumors. For this purpose, AlexNet used ESA's VGG16 and VGG19 models. In the analysis results, AlexNet, VGG16 and VGG19 models reached 89.95%, 94.65% and 94.82% accuracy rates, respectively. Rammurthy and Mahesh [9] proposed a Whale Harris Hawks optimization (WHHO) method based on deep convolutional neural network (DCNN). Their proposed method provided 81.6% accuracy in detecting the tumor. Nayak et al. [10] developed an algorithm based on DWT and DCNN. In studies, brain tumors were detected with 97% accuracy. Jakhar et al. [11] proposed multi-scale fractal feature network with segmentation. Their results achieved 94.66%, 94.42%, and 92.81% in accuracy, sensitivity, and specificity, respectively, using 5-fold cross-validation. Gupta et al. [12] performed a three-layer CNN to identify brain tumors. The sensitivity value of the analysis results was obtained as 0.90%.

Today, modern imaging techniques and genetic analysis biomarkers help detect brain tumors sensitively and quickly. However, research in this area still reveals the need to develop more effective screening methods and diagnostic techniques. A lot of data is obtained from images obtained with imaging techniques. It is very important to remove data that is not related to cancer and to understand the patterns that provide information about cancer. In this paper, a new hybrid cancer detection algorithm based on pre-trained features and kernel algorithm is proposed. Cancer (tumor) detection will be provided through images by using different machine learning algorithm architectures created from the extracted features.

This study aims to investigate the effectiveness of feature extraction and machine learning algorithms used in detecting brain tumors and to develop new and advanced methods for accurate diagnosis. For this reason, Deep learning and SVMs are utilized in this study, to create hybrid models. Deep learning architectures are distinguished from traditional artificial neural network models primarily by their capacity to generate a feature map of the data within their layers [13]. The feature maps that are produced by applying filters of various sizes and numbers to the data can be examined using these architectures. ResNets are one of the pre-trained deep neural networks used to extract information from MRI images. The idea was created using the ResNet architecture, which skips a layer between the inputs of the prior network and the outputs of the subsequent network to accomplish the combining process. It refers to this procedure as the residual block. This block's primary objective is to reduce data loss and ensure the model's dependability by processing the

prior values, even in the event that the intermediate values contain zero values. Additionally, it keeps the gradients that facilitate learning from disappearing [14-15-16-17]. Various image pre-processing methods are used in this work to extract MRI features. Image data augmentation is the process of creating changed reproductions of images that are in the same class as the original image. Data augmentation was performed before feature extraction to improve the model's capacity for generalization. Following the acquisition of MRI features tumor detection was carried out using an SVM-based kernel. SVM based kernel function was employed as a classifier. Kernel functions transform the data into a higher-dimensional space, making it linearly separable. Because of the nonlinear structure of features, kernel structures are used. By utilizing kernel structures, SVMs lessen the possibility of overfitting the individual model [17-18]. So, it is possible to produce forecasts that are more consistent and trustworthy. As a result, the model performs better overall, and the error rate decreases. The suggested algorithm's main framework is displayed in Figure 1.



**Figure 1.** The proposed main structure.

### 3. Materials and Methods

#### 3.1. Convolution neural network

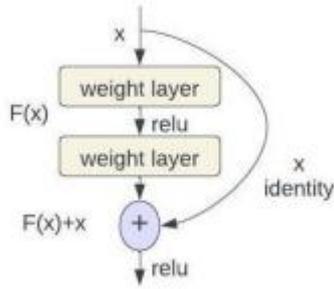
The convolutional neural network is designed using a multi-layer network topology. Yann LeCun created this network architecture, known as LeNet, in 1998 [19]. It filters the photographs' pixel matrices and automatically extracts distinctive features. Convolution layers serve as the foundation for deep neural networks. The convolution layer improves feature representation by integrating the convolution, activation function, and pooling phases to extract low-dimensional features from high-dimensional inputs [20]. The classification layer matches low-dimensional features and classes and is frequently a fully connected neural network. These layers allow CNN to recognize trends and address related issues [14-15-16-17-21]. CNNs allow for the extraction of features from high-dimensional data. The pretrained deep neural networks Residual Neural Network (ResNet) is described as follows:

*Residual Neural Network:*

Residual Neural Network (ResNet) is a CNN model that performed strongly in The ImageNet Large Scale Visual Recognition Challenge in 2015. The formula for the residual post-connection layer output is presented below [22].

$$H(x) = f(wx + b) + x \quad (1)$$

where  $x$  is the layer's input,  $f$  is the activation function, and  $b$  is the bias. Figure 2 depicts the fundamental blocks of residual learning.

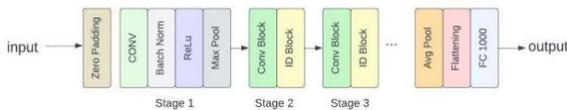


**Figure 2.** Residual learning: a building block [24]

$$F(x) = H(x) - x \tag{2}$$

$$y = F(x) + x \tag{3}$$

As seen in Equations 2 and 3,  $F(x)$  and  $y$  represent the residual function and output, respectively. The residual connection is a straight line that connects the layer input ( $x$ ) to the layer output, bypassing one or more layers. Inputs in residual blocks can spread more quickly across layers. Additionally, learning comparable mappings in the network is made easier by bypassing levels. CNN model performance starts to decline as the number of layers rises and the network gets deeper. CNN models based on blocks have now been created to address this issue. Figure 3 displays ResNet's general design [22]:



**Figure 3.** ResNet Structure. [24]

There are several pretrained networks available, including ResNet18, ResNet50, and ResNet101. To prevent overfitting, ResNet is created up of multiple residual blocks that are built on top of one another. According to He et al., all blocks use convolution and pooling layers to process 224x224 pixel input pictures. ResNet18 includes 18 layers, while ResNet50 and ResNet101 CNN models have 50 and 101 layers, respectively [22-23-24].

### 3.2. Support vector machine

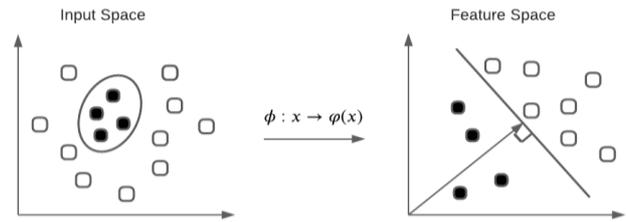
Support vector machine is one of the supervised learning methods generally used in classification problems. The main purpose of SVM is to find an optimal hyperplane that separates the data into different classes. This hyperplane is selected to maximize the separation between the classes, called margin. The wider the margin, the better the separation of two or more classes. If samples enter the margin region, this is called soft margin. However, hard margin works if the data is linearly separable and is very sensitive to outliers. Let's assume that the sample set is given as follows:  $(x_i, y_i), i=1, 2, \dots, n, x_i \in \mathbb{R}^p$  and  $y_i \in \{-1, 1\}$ .  $x_i$ ,  $i$ th sample and  $y_i$  the class label, classification plane:

$$y = w^T \cdot x + b \tag{4}$$

As seen in Equation 4,  $w$  is the hyperplane coefficient vector, and  $b$  is the distance of the hyperplane from the origin. In there, a hyperplane is a plane that separates data in an  $n$ -dimensional space. SVMs map the input space  $X$  to a high-dimensional feature space. More flexible classifications are achieved with this mapping.

A nonlinear kernel is used to transform in higher dimensional space and is obtained the optimal hyperplane. In a nonlinear dataset, SVMs cannot plot a linear hyperplane, and kernel

structures are used, for the data into a higher dimensional space, making it linearly separable as seen in Figure 4.



**Figure 4.** Kernel Operator. [17]

In Figure 4,  $\phi$  shows the transformation related with different kernel structure, called linear, polynomial, radial basis function, quadratic, cubic, gaussian [17-18]. Some kernel functions and their parameters are listed in Table 1 [14-15-16].

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \tag{5}$$

**Table 1.** Kernel function and formula.

| Kernel Function | Formula                                                             |
|-----------------|---------------------------------------------------------------------|
| Linear          | $K(x_i, x_j) = \langle x_i, x_j \rangle$                            |
| RBF             | $K(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2)$                       |
| Gaussian        | $K(x_i, x_j) = \exp\left(-\frac{\ x_i - x_j\ ^2}{2\sigma^2}\right)$ |
| Polynomial      | $K(x_i, x_j) = (1 + \langle x_i, x_j \rangle)^1$                    |
| Quadratic       | $K(x_i, x_j) = (1 + \langle x_i, x_j \rangle)^2$                    |
| Cubic           | $K(x_i, x_j) = (1 + \langle x_i, x_j \rangle)^3$                    |

### 3.3. Evaluation metrics

It is possible to develop several models using various machine learning techniques using data. Metrics for evaluation are required to determine which of the developed models will produce superior outcomes. Through a comparison of the model's predictions and the actual outcomes, evaluation metrics offer insight into how well the model performed its forecast. This study employed the following evaluation criteria: Matthews Correlation Coefficient, F1 Score, False Positive Rate, sensitivity, specificity, accuracy, and precision. Every formula is provided in the table that follows [24-25-26-27-28].

**Table 2.** The formula of evaluation metrics.

| Evaluation Metrics | Formula                                             |
|--------------------|-----------------------------------------------------|
| Accuracy           | $(TP+TN)/(TP+TN+FP+FN)$                             |
| Precision          | $TP/(TP+FP)$                                        |
| Sensitivity        | $TP/(TP+FN)$                                        |
| F1_score           | $2*(Precision*Sensitivity)/(Precision+Sensitivity)$ |
| Dice Coefficient   | $2*TP/(2*TP+FN+FP)$                                 |

Samples that fall into the positive class and are accurately predicted are known as true positives -TP.

Samples that are accurately predicted to be in the negative class are known as true negatives -TN.

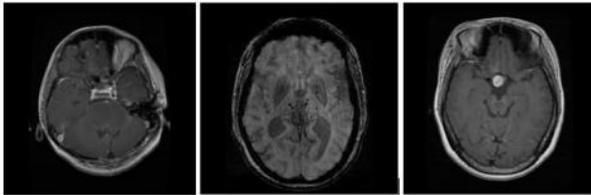
Samples from the negative class that were shown to be positive are known as false positives - FP.

Samples that are presented to be negative but belong to the positive class are known as false negatives - FN.

**4. Results**

**4.1. Dataset**

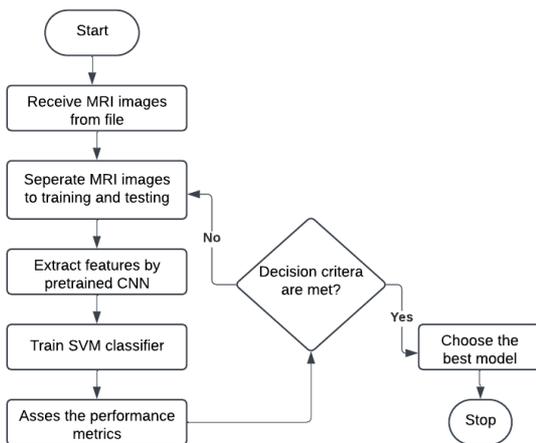
The dataset has three distinct MRI pattern classes: pituitary, glioma, and no tumor. There are 1621 MRI images for glioma, 1757 for pituitary, and 2000 for no tumor. There is more information available at [29]. Figure 5 shows some of the images.



**Figure 5.** MRI images from the dataset, glioma, pituitary and no tumor, respectively. [24]

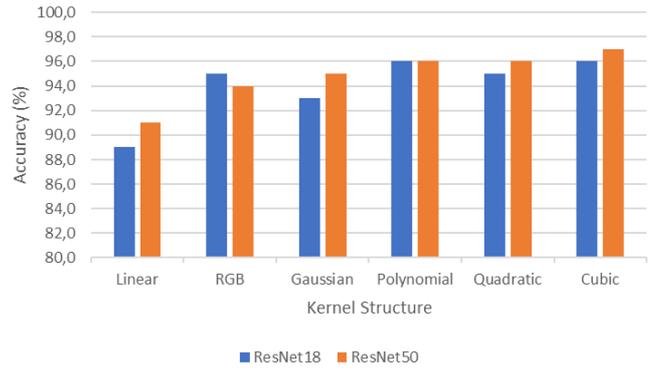
**4.2. Analysis**

ResNet18 and ResNet50 used RGB images with dimensions of  $224 \times 224$  pixels. ResNets are pre-trained models that use images from huge data sets. The pooling layer, which comes before the final classifier layer (totally connected) in these models, was employed for feature extraction. This layer received MRI pictures, which were used to obtain features. The identical procedure was followed for the Glioma, Pituitary, and No Tumor classes. All data sets are different sizes. To deal with imbalanced data distribution, the size of the smallest data collection was designated as the common size. The data were randomly divided into training and testing datasets with ratios of 0.8 and 0.2. The batch size was used, which represents the number of sub-samples transmitted to the network for each parameter update. The model parameters were obtained from the training dataset. Hyperparameters have a critical role in determining predictive model performance. The proposed algorithm in this work is shown in Figure 6.



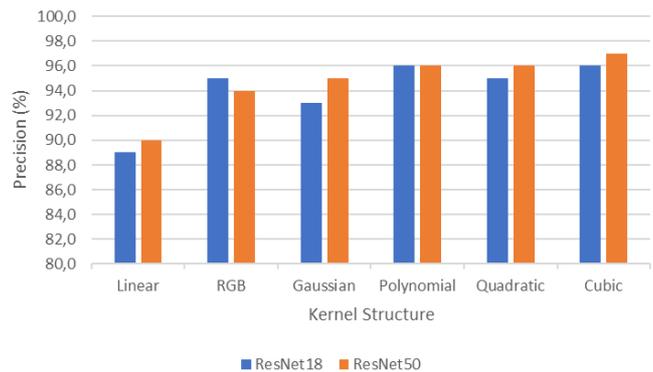
**Figure 6.** The proposed algorithm.

Analysis results are given between Figure 7 and 11, depending on accuracy, precision, sensitivity, F-measure, and dice coefficient as comparable to both feature extraction and kernel classifier.



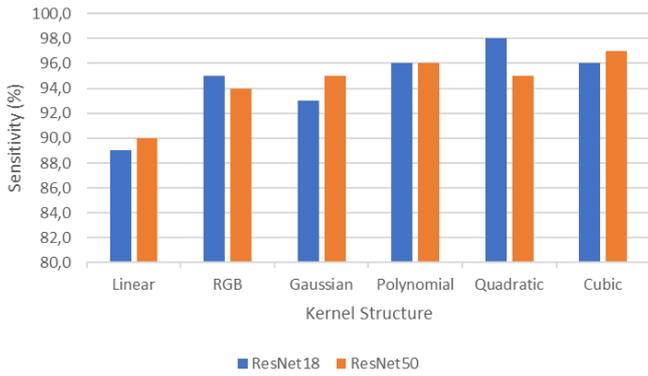
**Figure 7.** The comparative performance results based on accuracy.

Figure 7 compares the accuracy performance of ResNet18 and ResNet50 algorithms on different classifier models. ResNet50 shows a stronger performance by providing a higher accuracy rate than ResNet18 in all kernel structures, except RGB and Polynomial. Both algorithms achieved high accuracy rates in Polynomial, and Cubic kernel structures, which shows that they fit these kernel structures better. The highest accuracy rate was observed in the Cubic kernel structure based on ResNet50 pre-trained features. It is seen that Cubic kernel is the most suitable model for extracted features both ResNet18 and ResNet50. This analysis shows that the features extracted from ResNet50 provide better performance with more complex kernel structures and can generally be preferred.



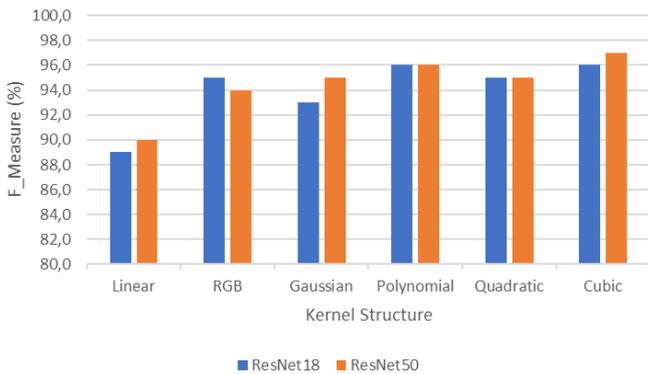
**Figure 8.** The comparative performance results based on precision.

Figure 8 compares the precision ratios of ResNet18 and ResNet50 algorithms on different classifier models. ResNet50 showed a generally more consistent performance by exhibiting a higher or similar precision ratio than ResNet18 in all kernel structures except the Linear kernel. In the Cubic kernel structure, both models reached high precision values. In the Linear kernel, the precision ratio of ResNet18 is lower compared to the other structures, which may indicate that the model makes more errors in this kernel. In the Polynomial kernel, both models obtained similar results. The highest precision ratio was observed in the Cubic kernel structure, which shows that this structure is especially suitable for features extracted from ResNet50 pre-trained model.



**Figure 9.** The comparative performance results based on sensitivity.

Figure 9 compares the sensitivity rates of ResNet18 and ResNet50 algorithms on different kernel structures. In the linear kernel, both pretrained algorithms performed low performance. In the RGB kernel structure, both algorithms gave similar results and the difference between them remained minimal. In the Gaussian kernel, ResNet50 achieved a lower sensitivity rate compared to ResNet18. In the Polynomial kernel structure, both methods gave the same results. But with Quadratic kernel structures, ResNet18 achieved a higher sensitivity rate compared to ResNet50, and showed that it is a more suitable option for these structures, unlike the Cubic kernel structure.

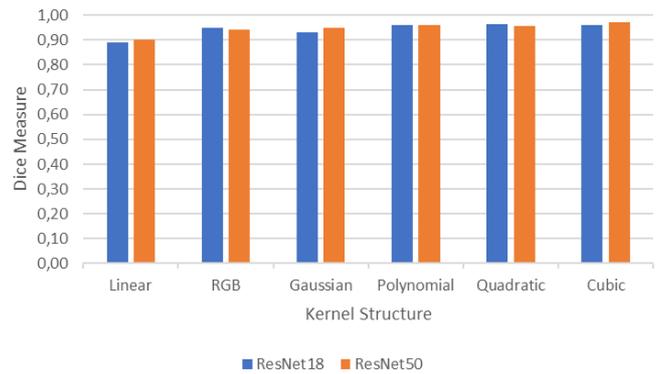


**Figure 10.** The comparative performance results based on F-measure.

Figure 10 compares the F-Measure performance obtained with different kernel structures for the ResNet18 and ResNet50 models. ResNet50 exhibits higher F-Measure values than ResNet18 in most kernel structures, with a significant difference especially in Gaussian and Cubic kernel structures. The lowest F-Measure value for ResNet18 and ResNet50 is observed in the Linear kernel structure. The highest performance is obtained in the RGB and Cubic kernel structure for ResNet18 and ResNet50 models, respectively.

Figure 11 compares Dice Coefficient performance with ResNet18 and ResNet50 models for different types of kernel structures. The graph shows that overall both models achieve a high Dice Coefficient (around 0.90), indicating a good prediction success. ResNet50 has a slightly better performance over ResNet18 in almost all kernel builds. In the Linear kernel structure, the performance for both models seems to be slightly lower compared to other kernel structures. In the "Polynomial," "Quadratic" and "Cubic" kernel structures, both models almost reached the maximum Dice Coefficient value. In general, it can be said that as the complexity of the kernel structure increases, the performance of the models remains consistently high, and the

features extracted from ResNet50 give slightly better classification results.



**Figure 11.** The comparative performance results based on Dice Coefficient.

As a result of this study, the comparative result of the best model is given in Table 3:

**Table 3.** The best classifier results.

| Kernel     | Pre-Trained CNN | Performance |
|------------|-----------------|-------------|
| Linear     | ResNet50        | 91 %        |
| RGB        | ResNet18        | 95 %        |
| Gaussian   | ResNet50        | 95 %        |
| Polynomial | ResNet18-50     | 96 %        |
| Quadratic  | ResNet18        | 98 %        |
| Cubic      | ResNet50        | 97 %        |

The tables above demonstrate that the ResNet18 (the 18-layer deep CNN) algorithm's features, when used to detect brain tumours, provide greater classifier success with quadratic kernel structure than other hybrid methods. Because this technique allows for feature extraction with fewer layers, it also produces a model that is less complicated.

**5. Conclusion**

Brain cancer is one of the most prevalent and lethal cancer forms. Early detection is critical in decreasing death rates from brain tumours. This paper proposes the ResNet18 and ResNet50 pretrained features and support vector machine hybrid model for early brain cancer detection. In this study, as a novelty to the studies in the literature, a hybrid algorithm is proposed that features were extracted with pre-trained based CNN, classification was made with SVM based different kernels, and in addition to classical classification metrics, the dice coefficient was also calculated. Malignant (glioma) tumours are successfully identified using the ResNet 18 and Quadratic kernel structure. Table 4 provides comparative findings from some research that have been published. The most impressive outcomes were highlighted in bold and italics.

**Table 4.** The performances of existing approaches in the literature.

| Autor's           | Method   | Performance |
|-------------------|----------|-------------|
| Vani et al. [5]   | SVM      | 82 %        |
| Mohsen et al. [6] | DWT, CNN | 93.94 %     |

**Table 4** (continued)

|                          |                       |             |
|--------------------------|-----------------------|-------------|
| Shahzadi et al. [7]      | AlexNet, VGGNet, LSTM | 84 %        |
| Swati et al. [8]         | AlexNet, VGG16, VGG19 | 94.82 %     |
| Rammurthy and Mahesh [9] | WHHO, DCNN            | 81.6 %      |
| Nayak et al. [10]        | DWT, DCNN             | 97 %        |
| Jakhar et al. [11]       | MSFFN                 | 94.66 %     |
| Gupta et al. [12]        | CNN                   | 90.0 %      |
| <i>This work</i>         | <i>ResNet, SVM</i>    | <i>98 %</i> |

MSFFN: Multi-scale fractal feature network

CNN: Convolutional Neural Network

DWT: Discrete Wavelet Transform

DCNN: Deep Convolutional Neural Network

SVM: Support vector machine

WHHO: Whale Harris Hawks optimization

The utilization of kernel structures in real-time data streams is the goal of future research. Thus, automatic model selection and hyperparameter setup optimization for support vector machine techniques will be investigated as a first step. In this manner, model combinations will be dynamically adjusted and optimized based on data features or changes over time. Additionally, it is assessed that the established and suggested procedures are adequately dependable and applicable to various imaging techniques and other disease classifications.

#### Ethics committee approval and conflict of interest statement

This article does not require ethics committee approval. This article has no conflicts of interest with any individual or institution.

#### Acknowledgments

This work was supported in part by the Scientific and Technological Research Council of TURKEY (TUBITAK) under grant No. 1059B141900679.

#### References

- Smith RA, Andrews KS, Brooks D, Fedewa SA, Manassaram-Baptiste D, Saslow D, et al. Cancer screening in the United States, 2017: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin* 2017;67(2):100-21. doi:10.3322/caac.21392.
- Lu G, Fakurnejad S, Martin BA, van den Berg NS, van Keulen S, Nishio N, et al. Predicting Therapeutic Antibody Delivery into Human Head and Neck Cancers. *Clin Cancer Res* 2020;26(11):2582-94. doi:10.1158/1078-0432.
- WHO. World Health Organization Report. <https://www.who.int/health-topics/cancer>; 2022.
- Kidwell CS, Hsia AW. Imaging of the Brain and Cerebral Vasculature in Patients with Suspected Stroke: Advantages and Disadvantages of CT and MRI. *Current Neurology and Neuroscience Reports* 2006;6(1):9-16. doi:10.1007/s11910-996-0003-1.
- Vani N, Sowmya A, Jayamma N. Brain Tumor Classification using Support Vector Machine. *International Research Journal of Engineering and Technology (IRJET)* 2017;4(7):792-6.
- Mohsen HM, El-Dahshan EA, El-Horbaty EM, Salem AM. Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal* 2017;3(1):68-71. doi:10.1016/j.fcij.2017.12.001.
- Shahzadi I, Tang TB, Meriadeau F, Quyyum A. CNN-LSTM: Cascaded framework for brain tumour classification. *IEEE EMBS Conference on Biomedical Engineering and Sciences* 2018:633-7.
- Swati ZNK, Zhao Q, Kabir M, Ali F, Ali Z, Ahmed S, et al. Brain Tumor Classification for MR Images using Transfer Learning and Fine-Tuning. *Computerized Medical Imaging and Graphics* 2019;75:34-46. doi:10.1016/j.compmedimag.2019.05.001.
- Rammurthy D, Mahesh PK. Whale Harris Hawks Optimization based Deep Learning Classifier for Brain Tumor Detection using MRI images. *Journal of King Saud University - Computer and Information Sciences* 2022;34(6):3259-72. doi:10.1016/j.jksuci.2020.08.006.
- Nayak DR, Padhy N, Mallick PK, Singh A. A deep autoencoder approach for detection of brain tumor images. *Computers and Electrical Engineering* 2022;102(2).
- Jakhar SP, Nandal A, Dhaka A, Alhudaif A, Polat K. Brain tumor detection with multi-scale fractal feature network and fractal residual learning. *Applied Soft Computing* 2024;153:1568-4946.
- Gupta BB, Gaurav A, Arya V. Deep CNN based brain tumor detection in intelligent systems. *International Journal of Intelligent Networks* 2024;5:30-7.
- Patterson J, Gibson A. *Deep Learning A Practitioner's Approach*. 1st ed. O'Reilly; 2017.
- Bishop CM. *Neural Network for Pattern Recognition*. Microsoft Research Cambridge; 1995.
- Bishop CM. *Pattern Recognition and Machine Learning*. Springer; 2006.
- Bishop CM. *Neural Networks for Pattern Recognition*. Oxford University Press; 2010.
- Ozer E. Early Diagnosis of Epileptic Seizures over EEG Signals using Deep Learning Approach. PhD Thesis. Mimar Sinan Fine Arts University, Institute of Science and Technology; 2023.
- Cortes C, Vapnik V. *Support-Vector Networks*. Kluwer Academic Publishers; 1995.
- LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 1998;86(11):2278-324.
- Wei D, Anurag B, Jianing W. *Deep Learning Essentials: Your Hands-on Guide to the Fundamentals of Deep Learning and Neural Network Modeling*. Packt Publishing; 2018.
- Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. Springer; 2009.
- He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA; 2016, p. 770-8.
- Dong Y, Zhang H, Wang C, Wang Y. Fine-Grained Ship Classification based on Deep Residual Learning for High-Resolution SAR Images. *Remote Sens Lett* 2019;10(11):1095-104.
- Özer E. Brain Tumor Detection using Deep CNNs and Ensemble Algorithms over MRI Images. *Computer Science* 2024;9(2):142-50. doi:10.53070/bbd.1455902.
- Van Rijsbergen CJ. *Information Retrieval*. London: Butterworths; 1979.
- Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters* 2006;27(8):861-74. doi:10.1016/j.patrec.2005.10.010.
- Powers D. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies* 2011;2(1):37-63. doi:10.48550/arXiv.2010.16061.
- Özer E. Detection of Brain Tumor using Boosting Algorithms based on Feature Selection. *Researcher* 2024;4(2):130-40.
- Nickparvar M. Brain Tumor MRI Dataset, Kaggle. doi:10.34740/KAGGLE/DSV/2645886; 2021.