World Neurosurgery: X 20 (2023) 100231

Contents lists available at ScienceDirect

World Neurosurgery: X

journal homepage: www.journals.elsevier.com/world-neurosurgery-x

Machine learning-based classification of pineal germinoma from magnetic resonance imaging

Suchada Supbumrung, Anukoon Kaewborisutsakul, Thara Tunthanathip *

Division of Neurosurgery, Department of Surgery, Faculty of Medicine, Prince of Songkla University, Songkhla, Thailand

ARTICLE INFO

Keywords:
Image classification
Machine learning
Histogram of oriented gradients
Local binary pattern
Pineal tumor

ABSTRACT

Introduction: Surgical approaches for tissue diagnosis of pineal tumors have been associated with morbidity and mortality. The classification of images by machine learning (ML) may assist physicians in determining the extent of resection and treatment plans for a specific patient. Therefore, the present study aimed to evaluate the diagnostic performances of the ML-based models for distinguishing between pure and non-germinoma of the pineal area. In addition, the secondary objective was to compare diagnostic performances among feature extraction methods.

Methods: This is a retrospective cohort study of patients diagnosed with pineal tumors. We used the RGB feature extraction, histogram of oriented gradients (HOG), and local binary pattern methods from magnetic resonance imaging (MRI) scans; therefore, we trained an ML model from various algorithms to classify pineal germinoma. Diagnostic performances were calculated from a test dataset with several diagnostic indices.

Results: MRI scans from 38 patients with pineal tumors were collected and extracted features. As a result, the k-nearest neighbors (KNN) algorithm with HOG had the highest sensitivity of 0.81 (95% CI 0.78–0.84), while the random forest (RF) algorithm with HOG had the highest sensitivity of 0.82 (95% CI 0.79–0.85). Moreover, the KNN model with HOG had the highest AUC, at 0.845. Additionally, the AUCs of the artificial neural network and RF algorithms with HOG were 0.770 and 0.713, respectively.

Conclusions: The classification of images using ML is a viable way for developing a diagnostic tool to differentiate between germinoma and non-germinoma that will aid neurosurgeons in treatment planning in the future.

1. Introduction

Brain tumors in the pineal region are infrequent, making up around 0.4% to 1% of all intracranial malignancies. Pineal tumors are also found in 10% of pediatric patients, but only 1%–4% of adult patients.[1,2] Various types of tumors, including such as germ cell tumors, pineal parenchymal tumors, gliomas, other tumors, and intracranial cysts, can occur in this region due to its diverse histological characteristics.[3] Therefore, tissue samples are required by surgical management for histological diagnosis.[4] Nonetheless, the surgical technique in the pineal area is still challenging, with morbidity reaching almost 20% in some studies and related mortality continuing at approximately 2%.[5] From prior studies, postoperative complications of pineal tumors treated with various microsurgical approaches included extraocular muscle dysfunction in 16–31% of cases, postoperative hemorrhage in 6–11% of cases, altered mental status in 5% of cases, and hemianopsia in 4% of cases.[6,7] In addition, endoscopic and stereotactic intraventricular biopsies are the two most used minimally invasive techniques for tissue biopsy in the pineal area. Postoperative complications associated with minimally invasive procedures include a new neurological deficiency ranging from 2.9% to 8.2%, postoperative

Abbreviations: AFP, α-fetoprotein; AI, Artificial intelligence; ANN, Artificial neural network; AUC, Area under the receiver operating characteristic curve; CCDSS, Computerized clinical decision support systems; CNN, Convolutional neural network; DT, Decision tree; FLAIR, Fluid-attenuated inversion recovery; HCG, Human chorionic gonadotropin; HOG, Histogram of oriented gradients; KNN, k-nearest neighbors; LBP, Local binary pattern; ML, Machine learning; mIU/ml, Milli-international units per milliliter; MRI, Magnetic resonance imaging; NB, Naive bayes; ng/ml, Nanograms per milliliter; NPV, Negative predictive value; PPV, Positive predictive value; RF, Random forest; SD, Standard deviation; SVM, Support vector machine; T1, T1-weighted; T1-Gd, T1-weighted gadolinium-enhanced; T2, T2-weighted; URL, Uniform resource locator; 95%CI, 95% confidence interval.

* Corresponding author. Division of Neurosurgery, Department of Surgery, Faculty of Medicine, Prince of Songkla University, Songkhla, 90110, Thailand.
E-mail address: tsus4@hotmail.com (T. Tunthanathip).

https://doi.org/10.1016/j.wnsx.2023.100231
Received 24 February 2023; Received in revised form 12 May 2023; Accepted 12 June 2023

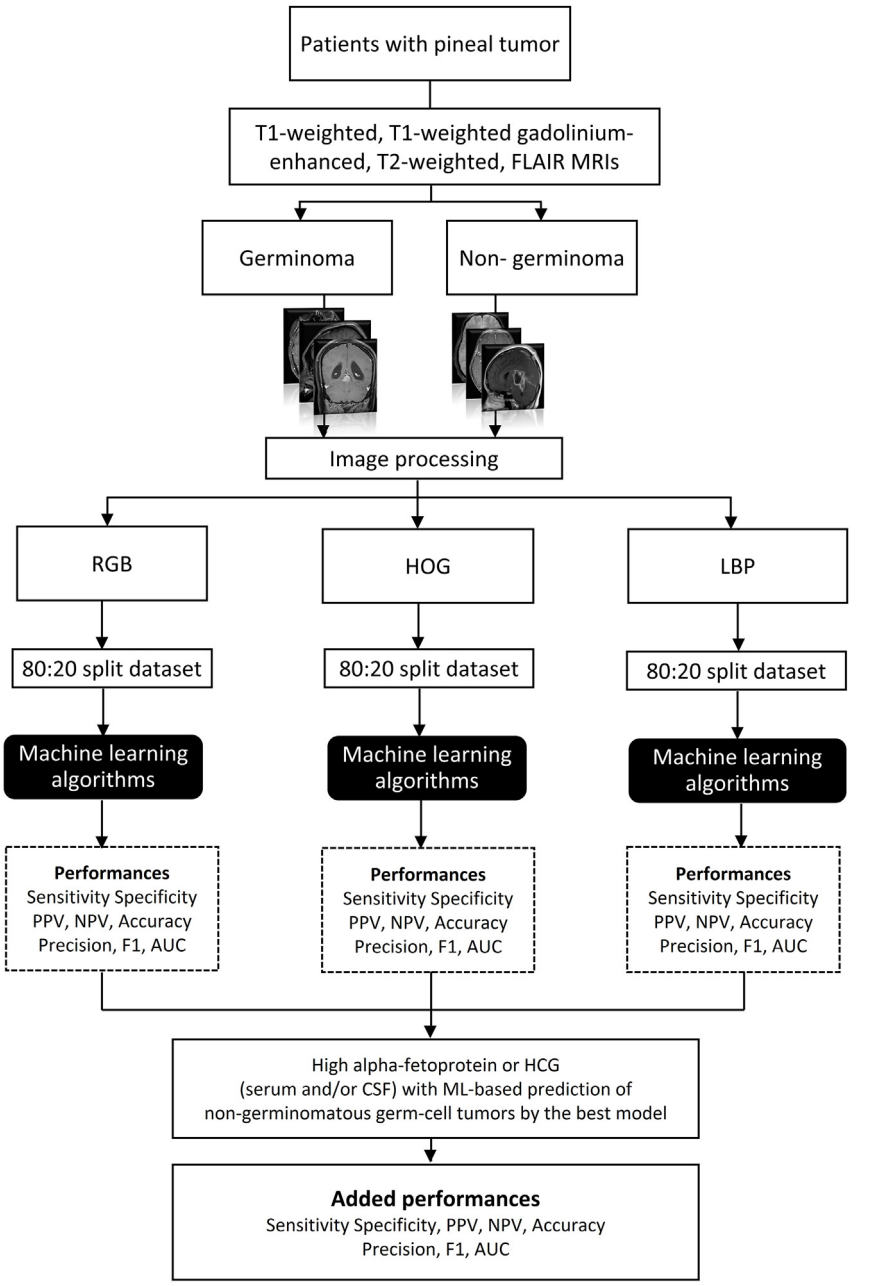

Fig. 1. Study flow of image classification for pineal tumors.

bleeding ranging from 4.4% to 19%, and perioperative death of 0%–1.9%.[8–11]

Because germinoma is a radiosensitive tumor; therefore, previous studies performed advanced imaging techniques to distinguish germinoma from other malignancies, such as advanced magnetic resonance imaging (MRI) scans or texture analysis, and machine learning (ML). While Ye et al extracted radiomics features with ML model training from MRI and reported that the highest area under the receiver operating characteristic curve (AUC) of the radiomics-only model was 0.80 (confidence interval (CI) 95%: 0.74–0.86),[12], Chen et al used texture analysis from MRI scans with ML algorithm for distinguish between germinoma and craniopharyngioma in the suprasellar region and the AUC was 0.91.[13]

The image classification by artificial intelligence (AI) may influence

neurosurgeons' and clinicians' decisions regarding the extent of resection and treatment options due to complications arising from getting tissue samples from the pineal area following surgery. Furthermore, there is a lack of evidence of the ML-based diagnostic performance from a review of the literature. Therefore, the present study aimed to evaluate the diagnostic performances of the ML-based models for distinguishing between pure and non-germinoma of the pineal area. In addition, the secondary objective was to compare diagnostic performances among feature extraction methods.

2. Methods

2.1. Study design and study population

This is a retrospective cohort study of patients diagnosed with pineal tumors between January 2014 and December 2021. Following patients were excluded: Patients with unavailable T1-weighted (T1), T1-weighted gadolinium-enhanced (T1-Gd), T2-weighted (T2), and Fluid-attenuated inversion recovery (FLAIR) MRI images and those without an official histological diagnosis from a pathologist. In the present study, tumors were classified into two classes: germinoma and non-germinoma. In detail, non-germinoma included non-germinomatous germ-cell tumors and mixed germ-cell tumors which is a combination of germinoma with other malignancies. Furthermore, glioma, meningioma, and pineal cell tumors were excluded.

Demographic features and imaging characteristics were demonstrated by descriptive statistics. For continuous variables, mean and standard deviation (SD) were utilized, whereas percentages were used to characterize categorical data. The R version 4.4.0 program (R Foundation, Vienna, Austria) was utilized for statistical analysis.

The present study's workflow is demonstrated in Fig. 1. T1, T1-Gd, T2, and FLAIR MRI scans of pineal tumors were initially collected in the axial, coronal, and sagittal planes. As a result, the total number of images was randomly divided 80:20 for the ML model training and test diagnostic performance processes.

2.2. Feature extraction from MRI scans

Various methods of feature extraction were performed in the present study in the preprocessing stage, as follows: RGB feature extraction, Histogram of Oriented Gradients (HOG), and local binary pattern (LBP).

In the RGB feature extraction, the grayscale MRI images were resized to 128 × 128 pixels and converted into RGB-colored images. Therefore, feature arrays of the three channels were converted to the 1-dimensional data frame. In feature engineering for images, HOG has been employed as a feature descriptor.[14] HOG descriptors were used for feature extraction after resizing images to 128 × 128 pixels. Each HOG grid was 8 × 8 pixels in size and put into a 9-bin histogram. After HOG feature creation, feature arrays were converted to data frames before ML model training. Furthermore, LBP is a well-known and commonly utilized feature for classification issues,[14], and it was used as a feature descriptor in this research. A histogram was used to summarize the probability of texture pattern, and LBP values were converted to data frames before ML model development.

2.3. Machine learning model development

In the training processes, the data frames of extracted features from three feature extraction methods were used for training the model by various supervised algorithms. The supervised ML algorithms were performed for training the model with ten-fold cross-validation as follows: support vector machine (SVM), logistic regression (LR), naive Bayes (NB), k-nearest neighbors (KNN) with $k = 3$, decision tree (DT), random forest (RF) and artificial neural network (ANN). The ML was performed using Python version 3.8.7 (Python Software Foundation) with the "scikit-learn" package (scikit-learn developers). Therefore, the ML models were tested by unseen MRI scans from the test dataset and diagnostic performances were estimated as follows: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, F1 score, and AUC.

Moreover, the ROC curves among various ML algorithms were compared in a single plot and the model's discrimination was evaluated by the AUC. Acceptable discrimination would correspond to an AUC of 0.7, whereas high and outstanding discrimination would correspond to an AUC of 0.8 and 0.9, respectively.[15,16]

According to the international Delphi agreement on the management of intracranial germ-cell tumors, both α-fetoprotein (AFP) and human chorionic gonadotropin (HCG) from serum and cerebrospinal fluid are parameters associated with surgical decision-making. Patients with high

Table 1
Clinical characteristics of patients with pineal tumors.

Factor	Germinoma (n = 22)	Non-germinoma (n = 16)
Gender		
Male	21 (95.5)	15 (93.8)
Female	1 (4.5)	1 (6.3)
Mean age-year (SD)	17.76 (6.60)	11.43 (7.22)
Age group-year		
≤18	14 (63.6)	14 (87.5)
>18	8 (36.4)	2 (12.5)
Number of tumor		
Single	16 (72.7)	15 (93.8)
Multiple	6 (27.3)	1 (6.3)
Suprasellar/pineal bifocal tumor	4 (18.2)	1 (6.3)
Mean size of tumor-cm (SD)	3.44 (1.43)	3.47 (1.55)
Preoperative leptomeningeal dissemination	4 (18.2)	3 (18.8)
Preoperative hydrocephalus	13 (59.1)	11 (68.8)
Mean biomarker level		
Serum alpha fetoprotein-ng/ml	1.88 (0.90)	234.17 (383.92)
CSF alpha fetoprotein-ng/ml	2.26 (1.33)	41.98 (70.07)
Serum beta HCG-mIU/ml	3.24 (3.12)	627.23 (222.91)
CSF beta HCG-mIU/ml	11.64 (19.77)	846.51 (2120.69)
Surgical approach		
Endoscopic approach	8 (36.4)	2 (12.5)
Microscopic approach	14 (63.6)	14 (87.5)
Extent of resection		
Biopsy	13 (59.1)	5 (31.3)
Tumor resection	9 (40.9)	11 (68.8)
Histological diagnosis		
Pure germinoma	22 (46.8)	–
Mature teratoma	–	4 (25.0)
Yolk sac tumor	–	3 (18.8)
Immature teratoma	–	2 (12.5)
Mixed germinoma with mature teratoma	–	2 (12.5)
Mixed germinoma with immature teratoma, choriocarcinoma, and embryonal carcinoma	–	1 (6.3)
Mixed germinoma with choriocarcinoma	–	1 (6.3)
Mixed germinoma with yolk sac tumor	–	1 (6.3)
Mixed mature teratoma with embryonal carcinoma	–	1 (6.3)
Mixed mature teratoma with choriocarcinoma	–	1 (6.3)

Abbreviation: CSF; cerebrospinal fluid, mIU/ml: milli-international units per milliliter, ng/ml: Nanograms per milliliter, SD; Standard deviation.

levels of biomarkers are thought to be malignant non-germinomatous components of tumors.[17] Non-germinomatous germ-cell tumors are the most likely diagnosis in these patients. Therefore, we combined the ML technique with biomarkers to improve predictive performance. The ML algorithms with the greatest AUC were chosen to evaluate the added value of predicting performance when combined with biomarker profiles. In detail, the ML-model prediction would be modified to non-germinomatous germ-cell tumors when -fetoprotein or HCG levels exceeded normal ranges. The normal value of AFP at our institutes was less than 7 ng per milliliter (ng/ml), and the normal value of HCG was less than 10 milli-international units per milliliter (mIU/ml).

2.4. Ethical considerations

The study was approved by a human research ethics committee (REC.65-431-10-1). The present study did not require patients' informed consent because it was a retrospective analysis. However, the patient identification numbers were encoded before the feature extraction processes.

3. Results

3.1. Baseline characteristics of pineal patients

As shown in Table 1, thirty-eight patients in the present study had pineal tumors, 22 of them had germinoma and 16 of them had non-germinoma. The average ages of germinoma and non-germinoma were 17.76 (standard deviation (SD) 6.60) and 11.43 years (SD 7.22), respectively. In both groups, there were more men than women. The germinoma group had an average tumor size of 3.44 cm (SD 1.43), while the non-germinoma group had a mean maximum tumor diameter of 3.47 cm (SD 1.55).

Bifocal lesions at suprasellar and pineal regions in germinoma were observed in 18.2%, with preoperative leptomeningeal dissemination in 18.2% of germinoma cases. Preoperative hydrocephalus was found in 63.2% of those patients in the present study. In germinoma, the microscopic operation was performed in 63.6%, while 87.5% of the non-germinoma group underwent surgery by microscopic surgery. Additionally, there were three patients with non-pure germinoma who had exceptionally high levels of AFP and HCG. In detail, a case of mixed mature teratoma with choriocarcinoma had serum HCG 8046.0 mIU/ml and CSF HCG 6247.0 mIU/ml, while mixed germinoma with choriocarcinoma and mixed germinoma with yolk sac tumor cases had CSF HCG 5428.00 mIU/ml and had serum AFP 918.5 ng/ml, respectively.

3.2. ML model development and validation

As a result, the total images comprised 6056 images, of which 3073 MRI scans were germinoma and 2983 were non-germinoma. Following feature extraction among various methods, the data frames were randomly split for the train-test ML model processes.

Diagnostic performances of ML algorithms are demonstrated according to the feature extraction methods of images in Table 2. In the RGB feature extraction, KNN algorithms had the highest sensitivity with 0.81 (95% CI 0.78–0.84), while RF algorithms had the highest sensitivity with 0.82 (95% CI 0.79–0.85) in HOG methods. As demonstrated in Figs. 2 and 3, the KNN model of the HOG feature extraction received the greatest AUC at 0.845, with ANN algorithms having the highest AUC at

Table 2
Diagnostic performances for pure germinoma among models.

Model	Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)	Accuracy (95%CI)	F1 (95%CI)
RGB						
LR	0.59 (0.55–0.62)	0.63 (0.59–0.67)	0.63 (0.59–0.67)	0.59 (0.55–0.63)	0.61 (0.58–0.64)	0.61 (0.56–0.65)
SVM	0.81 (0.78–0.84)	0.69 (0.65–0.72)	0.73 (0.70–0.77)	0.78 (0.74–0.81)	0.75 (0.73–0.78)	0.77 (0.75–0.80)
RF	0.75 (0.72–0.79)	0.82 (0.79–0.85)	0.82 (0.79–0.85)	0.76 (0.72–0.79)	0.79 (0.76–0.81)	0.78 (0.74–0.82)
NB	0.50 (0.46–0.53)	0.66 (0.62–0.70)	0.61 (0.57–0.65)	0.55 (0.51–0.59)	0.58 (0.55–0.60)	0.55 (0.49–0.59)
KNN	0.81 (0.78–0.84)	0.66 (0.62–0.70)	0.71 (0.68–0.75)	0.76 (0.73–0.80)	0.73 (0.71–0.76)	0.76 (0.74–0.79)
ANN	0.63 (0.60–0.67)	0.75 (0.71–0.78)	0.73 (0.69–0.76)	0.66 (0.62–0.69)	0.69 (0.66–0.71)	0.68 (0.62–0.72)
DT	0.61 (0.57–0.65)	0.65 (0.62–0.69)	0.65 (0.61–0.69)	0.61 (0.57–0.65)	0.63 (0.60–0.66)	0.63 (0.58–0.67)
HOG						
LR	0.72 (0.69–0.76)	0.77 (0.74–0.80)	0.77 (0.74–0.80)	0.72 (0.69–0.76)	0.75 (0.72–0.77)	0.75 (0.70–0.78)
SVM	0.76 (0.72–0.79)	0.87 (0.84–0.90)	0.85 (0.82–0.88)	0.79 (0.76–0.82)	0.81 (0.79–0.84)	0.80 (0.75–0.83)
RF	0.82 (0.79–0.85)	0.69 (0.65–0.73)	0.74 (0.71–0.77)	0.79 (0.75–0.82)	0.76 (0.74–0.78)	0.78 (0.76–0.81)
NB	0.58 (0.54–0.62)	0.68 (0.64–0.71)	0.66 (0.62–0.69)	0.60 (0.56–0.64)	0.63 (0.60–0.65)	0.61 (0.56–0.65)
KNN	0.61 (0.57–0.65)	0.88 (0.86–0.91)	0.85 (0.82–0.88)	0.68 (0.65–0.71)	0.74 (0.72–0.77)	0.71 (0.64–0.75)
ANN	0.63 (0.59–0.67)	0.86 (0.84–0.89)	0.81 (0.78–0.85)	0.71 (0.68–0.74)	0.75 (0.72–0.77)	0.71 (0.64–0.75)
DT	0.64 (0.60–0.68)	0.62 (0.59–0.66)	0.62 (0.58–0.65)	0.65 (0.61–0.69)	0.63 (0.61–0.66)	0.63 (0.60–0.67)
LBP						
LR	0.52 (0.48–0.56)	0.70 (0.67–0.74)	0.65 (0.61–0.69)	0.58 (0.54–0.61)	0.61 (0.58–0.63)	0.58 (0.52–062)
SVM	0.52 (0.48–0.56)	0.74 (0.68–0.76)	0.66 (0.62–0.70)	0.58 (0.55–0.62)	0.61 (0.59–0.64)	0.58 (0.52–0.62)
RF	0.66 (0.62–0.70)	0.73 (0.69–0.76)	0.72 (0.68–0.76)	0.67 (0.63–0.70)	0.69 (0.67–0.72)	0.69 (0.64–0.72)
NB	0.50 (0.46–0.54)	0.72 (0.69–0.76)	0.66 (0.61–0.70)	0.58 (0.54–0.61)	0.61 (0.58–0.63)	0.57 (0.51–0.61)
KNN	0.68 (0.64–0.72)	0.68 (0.64–0.72)	0.69 (0.66–0.73)	0.67 (0.63–0.70)	0.68 (0.65–0.71)	0.69 (0.64–0.72)
ANN	0.64 (0.61–0.68)	0.60 (0.50–0.64)	0.63 (0.59–0.67)	0.61 (0.57–0.65)	0.62 (0.60–0.65)	0.64 (0.60–0.67)
DT	0.60 (0.56–0.64)	0.67 (0.63–0.70)	0.66 (0.62–0.69)	0.61 (0.57–0.65)	0.63 (0.60–0.66)	0.63 (0.58–0.66)
Combination of biomarkers and ML-based prediction for non-germinomatous germ-cell tumors						
KNN with HOG	0.87 (0.84–0.89)ikūk	0.89 (0.86–0.91)	0.89 (0.87–0.91)	0.86 (0.84–0.89)	0.88 (0.86–0.89)	0.88 (0.85–0.90)

Abbreviation: ANN = artificial neural network, CI = confidence interval, DT = Decision tree, HOG= Histogram of oriented gradients, KNN = k-Nearest Neighbors, LBP = local binary pattern, LR = logistic regression, NB = naïve Bayes, NPV = negative predictive value, PPV = positive predictive value, RF = random forest, RGB = RGB image processing, SVM = support vector machines.

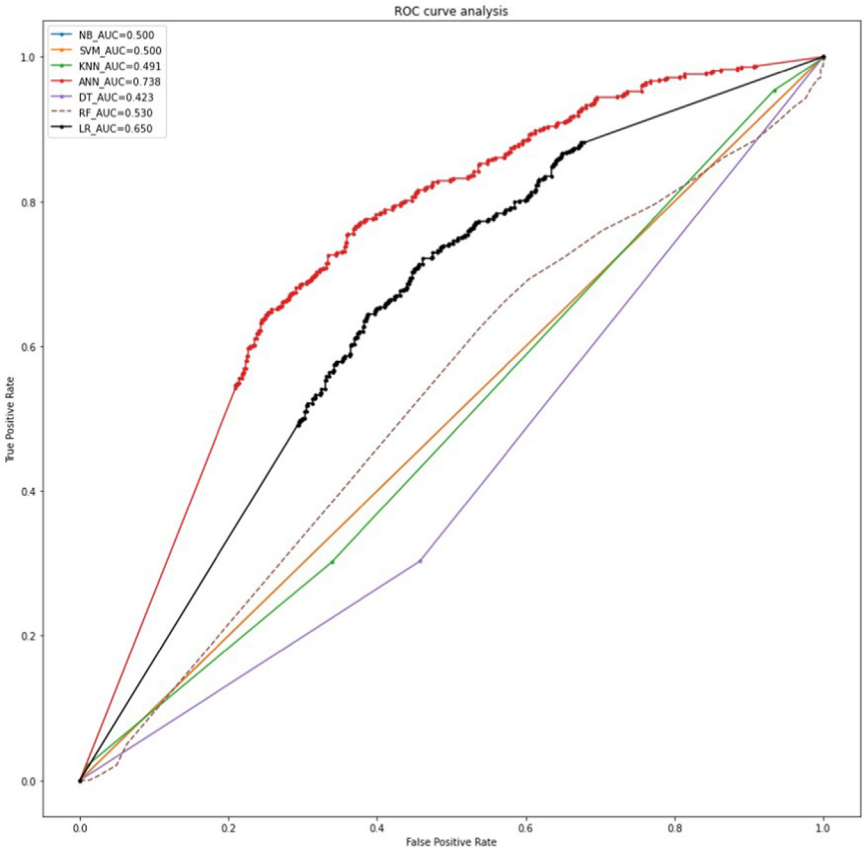

Fig. 2. The multiple ROC curves in a single plot according to RGB feature extraction with various ML algorithms. Abbreviation: ANN = artificial neural network, AUC = area under the receiver operating characteristic curve, DT = Decision tree, KNN = k-Nearest Neighbors, LR = logistic regression, NB = naïve Bayes, RF = random forest, SVM = support vector machines.

0.738 in the RBG feature extraction approach. Additionally, the highest AUC of the LBP feature extraction was the NB algorithm, as shown in Fig. 4.

Additionally, biomarkers for non-germinomatous germ-cell tumors were taken into prediction with the KNN model with HOG that had the highest AUC. As a result, the combined approach improved sensitivity from 0.61 (95%CI 0.57–0.65) to 0.87 (0.84–0.89) and enhanced the value of AUC from 0.845 to 0.877, as shown in Fig. 5.

4. Discussion

Germinomas are extraordinarily radiosensitive, and chemotherapy is used to treat them.[17–19] Image classification should be helpful for neurosurgeons to plan treatment strategies for each individual. As a result, ML-based classification from preoperative MRI scans had diagnostic prediction at a high level by AUC value, particularly KNN with HOG. However, the RF model with HOG may be a choice for application in real-world practice as the screening tool. Because the screening tool in medicine should have high sensitivity performance that there were few false negative results, and thus fewer cases of germinoma were missed.[24] We might employ the KNN model with HOG following the RF model with HOG for real-world applications because the model had a high specificity and AUC that was effective for confirming non-germinoma patients who prefer surgery to get tissue specimens.[20]

These results are in concordance with previous studies that ML had the role of image classification for germinoma. Chen et al used texture analysis with ML to discriminate between germinoma and craniopharyngioma in the suprasellar region, with T1-Gd scans providing the best AUC of validation.[13] Moreover, Ye et al report that the highest AUC of the radiomics-only model for pineal germinoma prediction was 0.80 (95%CI 0.74–0.86) when radiomics characteristics were performed from MRI scans using ML models.[12]

HOG and LBP have been used for image classification in intracranial tumors and tumor grading from the literature review.[21,22] Chen et al used HOG with an SVM algorithm to detect brain tumors from MRI scans and the grade of glioma. As a result, sensitivity, specificity, and AUC for intracranial tumor detection were 0.884, 0.805, and 0.921, respectively. Moreover, sensitivity, specificity, and AUC for glioma grading were 0.837, 0.687, and 0.806, respectively.[21] Kaplan et al utilized LBP with various ML algorithms for brain tumor classification and reported that accuracy scores for meningioma, glioma, and pituitary adenoma were 0.932, 0.905, and 0.955, respectively by LBP with KNN.[22] Brain edema or tumor necrosis areas can be observed in glioma and meningioma, which are useful for HOG and LBP feature extraction to distinguish these intracranial tumors, whereas the pineal region's complex structures, including the deep venous system, ventricular structures, cerebral aqueduct of Sylvius, and tectum, pose a classification challenge.

Furthermore, AFP and HCG are the potential biomarkers for

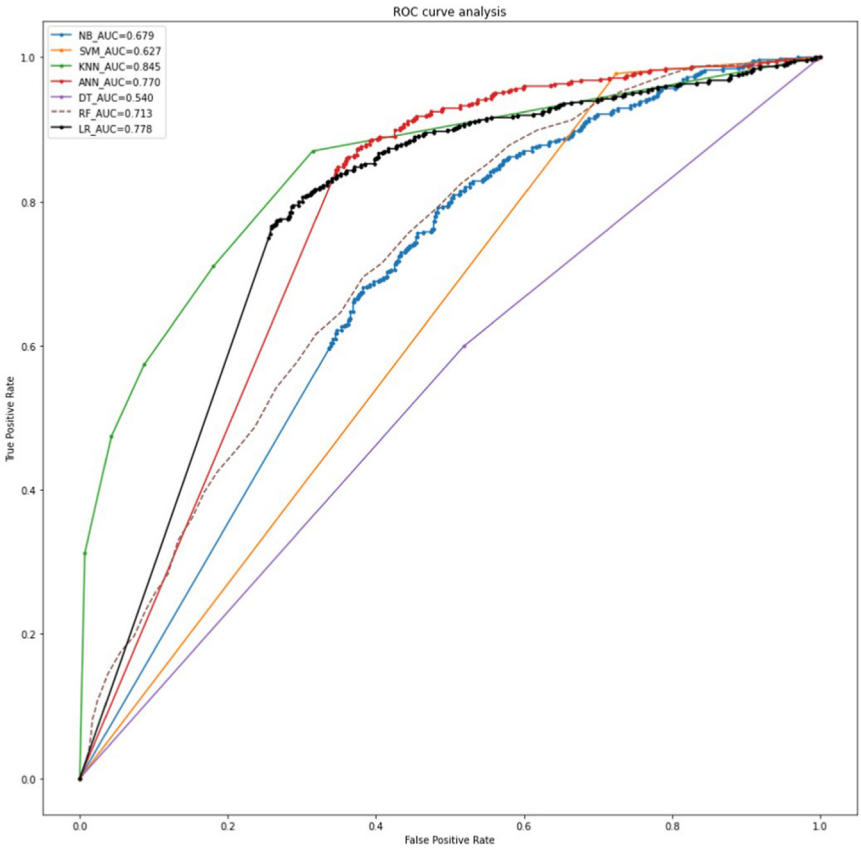

Fig. 3. The multiple ROC curves according to HOG feature extraction with various ML algorithms. Abbreviation: ANN = artificial neural network, AUC = area under the receiver operating characteristic curve, DT = Decision tree, KNN = k-Nearest Neighbors, LR = logistic regression, NB = naïve Bayes, RF = random forest, SVM = support vector machines.

malignant non-germinomatous components of tumors, such as choriocarcinoma and yolk sac tumor.[17] When the combined approach between the ML-based model and biomarkers, we observed the added sensitivity, NPV, and AUC that is useful for implication in general practice as the screening tool.[23]

Image classification based on AI may support a neurosurgeon with preoperative planning such as the extent of resection strategies or a re-operation decision in a patient who could not be diagnosed from a prior operation. From the literature review, Choi et al proposed treatment strategies for pineal tumors. In detail, surgery after a trial of chemotherapy and/or radiotherapy is suggested for a pineal tumor with negative tumor markers and suspected germinoma based on MRI.[19] Hence, AI-based diagnosis may assist physicians in proposing treatment options in the future.

Clinicians use evidence-based medicine, guidelines, and the results of clinical trials to make decisions. Thus, AI has evolved into a screening tool that can assist physicians in making a decision. It is a challenge to deploy machine learning models in general practice, so the ML-based tool needs to be kept as straightforward and user-friendly as possible in the form of a computerized clinical decision support system (CCDSS).[24] For instance, Tunthanathip et al predicted an intracranial hemorrhage in a child who had had a traumatic brain injury by employing an ML-based online application that was presented in CCDSS format. This was done to reduce unnecessary expenditures and prevent over radiation exposure.[25]

To the authors' knowledge, the present study was the first paper that used HOG and LBP feature extractions with ML for image classification between pineal germinoma and non-germinoma. However, the limitations of the present study should be noted. The CCDSS may require external validation before implementation; therefore, Multicenter studies should offer additional images of the precise tumor location for training the ML model and improving this issue.[26] Additionally, the ML-based model with feature extraction requires the expertise of specialists for the region of interest selection. Therefore, the deep learning method may be the alternative way for the classification of pineal germinoma because This method includes automatically the procedures of feature extraction and model construction via the convolutional neural network architecture.[27,28]

5. Conclusion

The classification of images using ML is a viable way for developing a diagnostic tool to differentiate between germinoma and non-germinoma that will aid neurosurgeons in treatment planning in the future.

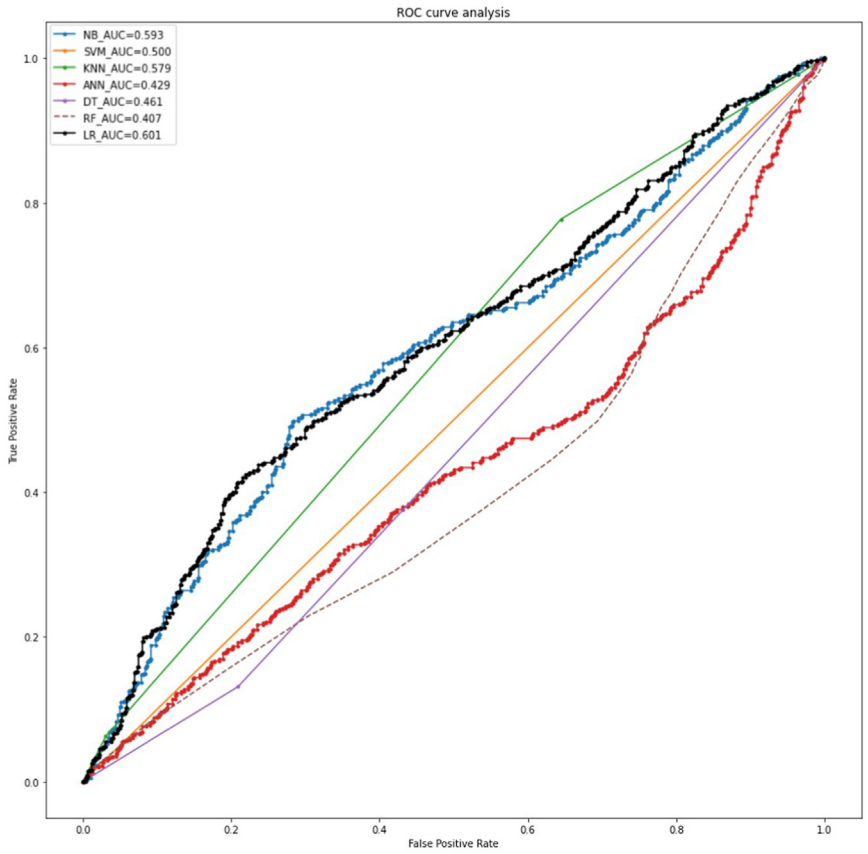

Fig. 4. The multiple ROC curves according to LBP feature extraction with various ML algorithms. Abbreviation: ANN = artificial neural network, AUC = area under the receiver operating characteristic curve, DT = Decision tree, KNN = k-Nearest Neighbors, LR = logistic regression, NB = naïve Bayes, RF = random forest, SVM = support vector machines.

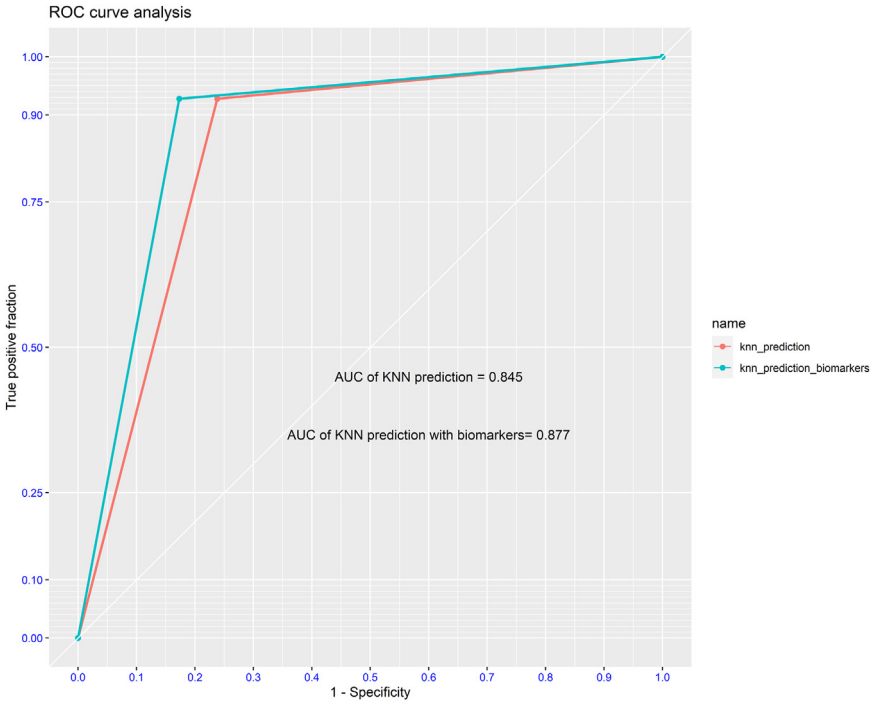

Fig. 5. The ROC curves of KNN algorithm with HOG feature extraction and combine KNN with HOG feature extraction and biomarkers. Abbreviation: AUC = area under the receiver operating characteristic curve, KNN = k-Nearest Neighbors.

Author contributions

TT conceived the study and designed the method. AK and SS supervised the completion of the data collection. TT undertook the recruitment of participating centers and patients and managed the data, including quality control. AK and SS provided statistical advice on the study design and analyzed the data, while AK drafted the manuscript, and all authors contributed substantially to its revision. TT takes responsibility for the paper as a whole.

Funding

None.

Research quality and ethics statement

The authors attest that it was determined this clinical investigation required Institutional Review Board/Ethics Committee review. The corresponding protocol/approval number is (REC.65-431-10-1).

CRediT authorship contribution statement

Suchada Supbumrung: Conceptualization, Data curation, Formal analysis. **Anukoon Kaewborisutsakul:** Conceptualization, Methodology, Visualization. **Thara Tunthanathip:** Conceptualization, Data curation, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Hirato J, Nakazato Y. Pathology of pineal region tumors. *J Neuro Oncol*. 2001;54(3):239–249. https://doi.org/10.1023/a:1012721723387.
2. Mottolese C, Szathmari A, Beuriat PA. Incidence of pineal tumours. A review of the literature. *Neurochirurgie*. 2015;61(2–3):65–69. https://doi.org/10.1016/j.neuchi.2014.01.005.
3. Fedorko S, Zweckberger K, Unterberg AW. Quality of life following surgical treatment of lesions within the pineal region. *J Neurosurg*. 2018;130(1):28–37. https://doi.org/10.3171/2017.7.JNS17260.
4. Pettorini BL, Al-Mahfoud R, Jenkinson MD, Avula S, Pizer B, Mallucci C. Surgical pathway and management of pineal region tumours in children. *Childs Nerv Syst*. 2013;29(3):433–439. https://doi.org/10.1007/s00381-012-1954-y.
5. Balossier A, Blond S, Touzet G, et al. Endoscopic versus stereotactic procedure for pineal tumour biopsies: comparative review of the literature and learning from a 25-year experience. *Neurochirurgie*. 2015;61(2–3):146–154. https://doi.org/10.1016/j.neuchi.2014.06.002.
6. Konovalov AN, Pitskhelauri DI. Principles of treatment of the pineal region tumors. *Surg Neurol*. 2003;59(4):250–268. https://doi.org/10.1016/s0090-3019(03)00080-6.
7. Bruce JN, Ogden AT. Surgical strategies for treating patients with pineal region tumors. *J Neuro Oncol*. 2004;69(1–3):221–236. https://doi.org/10.1023/b:neon.0000041885.09226.2d.
8. Song JH, Kong DS, Shin HJ. Feasibility of neuroendoscopic biopsy of pediatric brain tumors. *Childs Nerv Syst*. 2010;26(11):1593–1598. https://doi.org/10.1007/s00381-010-1143-9.
9. Pople IK, Athanasiou TC, Sandeman DR, Coakham HB. The role of endoscopic biopsy and third ventriculostomy in the management of pineal region tumours. *Br J Neurosurg*. 2001;15(4):305–311. https://doi.org/10.1080/02688690120072441.
10. Regis J, Bouillot P, Rouby-Volot F, Figarella-Branger D, Dufour H, Peragut JC. Pineal region tumors and the role of stereotactic biopsy: review of the mortality, morbidity, and diagnostic rates in 370 cases. *Neurosurgery*. 1996;39(5):907–914. https://doi.org/10.1097/00006123-199611000-00003.
11. Kreth FW, Schätz CR, Pagenstecher A, Faist M, Volk B, Ostertag CB. Stereotactic management of lesions of the pineal region. *Neurosurgery*. 1996;39(2):280–291. https://doi.org/10.1097/00006123-199608000-00010.
12. Ye N, Yang Q, Liu P, Chen Z, Li X. A comprehensive machine-learning model applied to MRI to classify germinomas of the pineal region. *Comput Biol Med*. 2022;152, 106366. https://doi.org/10.1016/j.compbiomed.2022.106366.
13. Chen B, Chen C, Zhang Y, , et alHuang Z, Wang H, Li R, Xu J. Differentiation between germinoma and craniopharyngioma using radiomics-based machine learning. *J Pers Med*. 2022;12(1):45. https://doi.org/10.3390/jpm12010045.
14. Jaruenpunyasak J, Duangsoithong R. Empirical analysis of feature reduction in deep learning and conventional methods for foot image classification. *IEEE Access*. 2021;9:53133–53145.
15. Tunthanathip T, Sae-Heng S, Oearsakul T, Kaewborisutsakul A, Taweesomboonyat C. Economic impact of a machine learning-based strategy for preparation of blood products in brain tumor surgery. *PLoS One*. 2022;17(7), e0270916. https://doi.org/10.1371/journal.pone.0270916.
16. Taweesomboonyat C, Tunthanathip T, Sae-Heng S, Oearsakul T. Diagnostic yield and complication of frameless stereotactic brain biopsy. *J Neurosci Rural Pract*. 2019;10(1):78–84. https://doi.org/10.4103/jnrp.jnrp_166_18.
17. Murray MJ, Bartels U, Nishikawa R, Fangusaro J, Matsutani M, Nicholson JC. Consensus on the management of intracranial germ-cell tumours. *Lancet Oncol*. 2015;16(9):e470–e477. https://doi.org/10.1016/S1470-2045(15)00244-2.
18. Choi JU, Kim DS, Chung SS, Kim TS. Treatment of germ cell tumors in the pineal region. *Childs Nerv Syst*. 1998 Jan-Feb;14(1–2):41–48. https://doi.org/10.1007/s003810050173.PMID:9548340.
19. Cho BK, Wang KC, Nam DH, , et alKim DG, Jung HW, Kim HJ, Han DH, Choi KS. Pineal tumors: experience with 48 cases over 10 years. *Childs Nerv Syst*. 1998;14(1–2):53–58. https://doi.org/10.1007/s003810050175.
20. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med*. 2013 Spring;4(2):627–635.
21. Chen T, Xiao F, Yu Z, Yuan M, Xu H, Lu L. Detection and grading of gliomas using a novel two-phase machine learning method based on MRI images. *Front Neurosci*. 2021;15, 650629. https://doi.org/10.3389/fnins.2021.650629. Published 2021 May 14.
22. Kaplan K, Kaya Y, Kuncan M, Ertunc HM. Brain tumor classification using modified local binary patterns (LBP) feature extraction methods. *Med Hypotheses*. 2020;139, 109696.
23. Maxim LD, Niebo R, Utell MJ. Screening tests: a review with examples. *Inhal Toxicol*. 2014;26(13):811–828. https://doi.org/10.3109/08958378.2014.955932.
24. Kwan JL, Lo L, Ferguson J, Goldberg H, Diaz-Martinez JP, Tomlinson G, et al. Computerized clinical decision support systems and absolute improvements in care: meta-analysis of controlled clinical trials. *BMJ*. 2020;370:m3216. https://doi.org/10.1136/bmj.m3216.
25. Tunthanathip T, Duangsuwan J, Wattanakitrungroj N, Tongman S, Phuenpathom N. Comparison of intracranial injury predictability between machine learning algorithms and the nomogram in pediatric traumatic brain injury. *Neurosurg Focus*. 2021;51(5):E7. https://doi.org/10.3171/2021.8.FOCUS2155.
26. Tunthanathip T, Oearsakul T. Application of machine learning to predict the outcome of pediatric traumatic brain injury. *Chin J Traumatol*. 2021;24(6):350–355. https://doi.org/10.1016/j.cjtee.2021.06.003. Epub 2021 Jun 8. PMID: 34284922; PMCID: PMC8606603.
27. Jacaruso LC. Accuracy improvement for Fully Convolutional Networks via selective augmentation with applications to electrocardiogram data. *Inform Med Unlocked*. 2021;26, 100729. https://doi.org/10.1016/j.imu.2021.100729.
28. Smirnov EA, Timoshenko DM, Andrianov SN. Comparison of regularization methods for ImageNet classification with deep convolutional neural networks. *AASRI Procedia*. 2014;6:89–99.