CENTIVE
Conference on Electrical Engineering, Informatics
, Industrial Technology, and Creative Media
ISBN : 978-602-53004-0-0

Automated Detection of Foot Tumor: A Machine Learning Approach Leveraging GLCM Texture Analysis

Asyafa Ditra Al Hauna[1*], Raphon Galuh Candraningtyas[2], and Yit Hong Choo[3]
[1]Telkom University
[2]Mie University
[3]Swinburne University
[1]Jl. DI Panjaitan No.128, Karangreja, Purwokerto Kidul, Kecamatan. Purwokerto Selatan, Kabupaten Banyumas, Jawa Tengah 53147, Indonesia
[2]1577 Kurimamachi Yachō, Tsu, Mie 514-0102, Japan
[3]John Street, Hawthorn VIC 3122, Australia
[*]alditra@student.telkomuniversity.ac.id

Abstract — Foot tumors are rare but diagnostically challenging due to overlapping symptoms with benign conditions. Automated image-based detection can aid early identification and reduce misdiagnosis. This study explores the use of GLCM-based feature extraction to classify foot magnetic resonance imaging, focusing on the presence or absence of tumors. The features were classified using logistic regression, decision tree, and random forest. Model performance was evaluated under a five-fold cross-validation framework with scaled features. Experimental results demonstrated strong classification performance, with all models achieving scores between 0.97 and 1.0 across defined metrics. Correlation analysis further revealed that homogeneity, energy, and angular second moment (ASM) had negative associations with the target, while other features showed positive correlations. These findings provide evidence that classical machine learning models, supported by feature engineering, are effective for the detection of foot tumors in absence and presence.

Keywords – foot tumor detection, MDCT, GLCM, machine learning, feature extraction

I. INTRODUCTION

The composition of the foot consists of complex structures surrounded by diverse tissues whereas each is capable of being the origin for the development of tumors [1]. Approximately a total of 80% of foot tumors are typically benign in pathology [2]. Based on a 20-year retrospective study conducted at a musculoskeletal tumor referral center, from a total of 2,660 tumors treated between 1986 and 2006, only 5.75% were in the foot and ankle [3]. Among the 153 cases of foot tumors, a significant proportion of nearly 40% are malignant, which is relatively high compared to other anatomical sites [3]. More than 90% of patients commonly present with localized swelling, and around 50% report experiencing pain [4]. However, specific clinical indicators such as the lesion size, pain, or prior trauma are often nonspecific and may not be reliable to indicate a tumor. Conspicuously, the malignant potential of foot tumors is frequently underestimated, as benign lesions can cause severe pain, whereas some malignant tumors may remain painless and progress slowly over several years [5], [6]. Given the rarity and diversity of foot tumors, coupled with the absence of pain, fractures, or other distinctive symptoms, a misdiagnosis can occur, often leading to delayed treatment and increased morbidity [4], [7]. Although plain radiographs are often used as an initial diagnostic tool, they provide limited information for tumor detection. multidetector computed tomography (MDCT), on the other hand, remains the gold standard for accurately characterizing soft tissue lesions [8]. However, interpreting MDCT scans requires considerable expertise and may be subject to diagnostic variability. Given these diagnostic challenges, researchers have increasingly turned to computational methods to support clinical decision-making. In this context, automated interpretation powered by deep learning offers the potential to act as an early screening tool, helping to flag suspicious tumors for further clinical evaluation.

Machine learning learning has long been used in the medical imaging field, among other applications

such as computer-aided diagnosis (CAD), radiomics, and medical image analysis [8]. The growth of medical imaging and Artificial Intelligence (AI) has created many possibilities in the field of healthcare. By combining these two fields, it has revolutionized various aspects of medical practices such as early diagnosis detection and accurate diagnosis to personalize treatment planning and improved patient outcomes. Automated image analysis itself in the medical field involved automatic parameterization of diagnostic images, aside from the manual physician evaluation that has been done. It produces more informative features than raw pixel intensities. In some cases, it achieves either comparable or outperform diagnostic accuracy than physicians, as it improves the sensitivity, specificity, and reliability through ROC-based analysis [9]. AI applications in medical fields have also been extensively explored for tumor detection. Such as how many Computer-Aided Diagnostic (CAD) systems have been developed to automatically detect abnormalities on the brain [10]–[12], Positron Emission Tomography (PET) imaging to classifying the stage of lung tumor [13], using AI to analyze, monitor, and interpret the complexity of tumor imaging, enabling real-time monitoring of tumorigenesis and biophysical properties through high-throughput imaging systems [14]. Despite significant advancements in applying AI to tumor detection and modeling, its applications in foot tumor detection are limited, thereby presenting a distinct gap within the broader landscape of tumor-focused medical imaging studies. This gap underscores the need to investigate alternative approaches that can work effectively even with limited datasets, such as combining machine learning with established feature engineering methods.

One of the main challenges in this domain is the limited availability of large-scale datasets, which highlights the importance of methods that can still extract meaningful information under such constraints. Machine learning methods have shown success in various imaging domains. However, for settings with limited datasets, handcrafted feature engineering remains valuable. Among the established approaches, texture-based methods such as the Gray Level Co-occurrence Matrix (GLCM) have demonstrated effectiveness in tumor characterization [15]. GLCM itself is a statistical method for analyzing texture in images [16]. It has been used in other studies about tumors detection such as lung tumors detection [17], for brain tumor classification [18], [19] , and other tumor-related tasks. However, research specifically addressing foot tumors remains scarce, leaving this area relatively underexplored compared to other tumor domains. Building on this gap, our work introduces a novel perspective by framing foot tumor analysis as a classification task. Whereas most tumor classification studies focus on distinguishing between different tumor types or variants against healthy tissue, we begin with the more fundamental step of separating tumor-present from tumor-absent cases. This framing provides a unique angle for our study and serves as a critical foundation for advancing AI-driven foot tumor diagnostics.

This study investigates the role of GLCM-based feature extraction in the classification of foot MDCT scans, focusing on present and absent tumor detection. The main contribution of this study as follows:

a) We explore how GLCM extracts discriminative features from foot tumor images.
b) We evaluate whether the engineered features contribute to strong model performance across different classifiers.
c) We assess the effectiveness of the resulting models in detecting the presence or absence of foot tumors.
d) We provide a foundation for future advancements by outlining how feature engineering approaches may be compared with filter-based machine learning and deep learning methods such as convolutional neural networks (CNNs).

By addressing these research questions, our work contributes novel insights into feature engineering for rare tumor detection and sets the groundwork for subsequent studies aimed at improving automated diagnostic tools for foot tumors.

II. LITERATURE RESEARCH

A. Foot Tumor: Clinical and Diagnostic Context

Although tumors of the foot and ankle represent only a small proportion of musculoskeletal neoplasms, their diagnosis poses unique challenges due to their rarity, overlapping presentation with pseudotumorous lesions, and the low clinical awareness among practitioners [20]. Epidemiological analyzes have also shown that while most of the soft tissue masses of the foot and ankle are benign, distinguishing malignant from benign lesions is not straightforward, as clinical indicators such as pain or duration of symptoms are unreliable. Instead, factors such as increasing patient age, larger lesion size ($\geq$4.0 cm), and male gender have been identified as significant predictors of malignancy, underscoring the complexity of risk stratification in routine clinical practice [21]. Conventional radiographs are often the first-line imaging tool in evaluating suspicious foot lesions, yet their ability to delineate soft-tissue tumors is limited, particularly when precise localization and tissue characterization are required. In contrast, MDCT has demonstrated high accuracy in assessing the anatomic extent and compartmental involvement of foot tumors, frequently allowing correct distinction between benign and malignant lesions and, in some cases, even suggesting specific benign diagnoses such as hemangiomas, ganglion cysts, or plantar fibromatosis [22]. MDCT is therefore considered the modality of choice for soft-tissue tumor assessment in the foot, as it provides clear delineation of tumor

margins and often identifies characteristic features of common benign and reactive processes, although histologic specificity remains limited [23]. Nevertheless, the inherent limitations of MDCT in achieving definitive histologic specificity, combined with the high cost and potential overuse of advanced imaging, underscore the need for automated diagnostic support tools that can assist clinicians in risk stratification and decision-making while optimizing resource utilization.

B. Medical Imaging Modalities in Tumor Diagnosis

As medical data continue to expand in both volume and diversity, radiology remains a central focus due to its high data density and diagnostic relevance [24]. Within this context, different imaging modalities play distinct roles in tumor diagnosis, each offering specific strengths and limitations. However, the interpretation of medical images has traditionally relied on the subjective judgment of radiologists, which introduces variability and potential limitations in consistency [25]. Recent advances point toward a transition to quantitative imaging (QI), where objective metrics, such as tumor growth rate, can enhance diagnostic precision and patient care. Comparative analyses of imaging modalities further emphasize that each technique carries distinct strengths and limitations, influencing its suitability for different diagnostic contexts [26]. Factors such as cost-effectiveness, accessibility, and interpretation variability also play a crucial role in determining their clinical application, while recent advances are reshaping their role in patient care [26]. Diagnostic imaging techniques (DITs) have become indispensable in modern medicine, offering noninvasive means to examine internal structures and monitor treatment response across multiple specialties. Among these, conventional radiography, ultrasonography, computed tomography (CT), and magnetic resonance imaging (MRI) represent the most widely applied modalities, each with unique principles, advantages, and challenges [27].

Radiography remains a valuable first-line modality, particularly effective in detecting and characterizing bone abnormalities due to its high spatial resolution and accessibility. However, its utility in tumor diagnosis is limited by difficulties in reliably detecting tumor boundaries, as edges often disappear, are obscured by surrounding anatomical structures such as ribs, or become buried in image noise, making interpretation highly challenging [28]. Ultrasonography provides real-time imaging that is cost-effective and widely available, with advanced techniques such as contrast-enhanced transrectal ultrasound (CE-TRUS) improving cancer detection rates through targeted biopsy. Nevertheless, its diagnostic performance remains constrained by false-positive findings—often caused by benign conditions such as hyperplasia or prostatitis—and false negatives, particularly in small or well-differentiated tumors, which limit its standalone reliability in comprehensive tumor evaluation [29]. Magnetic resonance imaging (MRI) has established itself as the gold standard for evaluating musculoskeletal bone and soft-tissue tumors, offering superior soft tissue contrast and the ability to characterize tumor morphology in great detail. Beyond conventional anatomical sequences, advances in functional and metabolic imaging further enhance its diagnostic value. However, its widespread application is challenged by the need to optimize protocols for diverse tissue types, heterogeneous tumor pathologies, and moving anatomical regions such as the chest wall, which complicate standardization across clinical practice [30]. Based on the three advantages and limitations, while radiography remains indispensable for bone assessment and ultrasonography offers real-time, accessible imaging with evolving enhancements, both modalities face critical limitations in reliably characterizing tumors. MDCT stands out by providing unparalleled soft tissue contrast, functional insights, and reproducible imaging features, making it the most suitable foundation for automated tumor analysis. Therefore, the strengths of MDCT in delivering detailed, quantitative, and consistent imaging data not only establish it as the clinical gold standard but also justify its central role in advancing machine learning–based diagnostic systems for tumor evaluation.

C. Feature Extraction in Tumor Image Analysis

Handcrafted feature extraction methods remain valuable in medical image analysis, particularly when datasets are relatively small and interpretability is important. Among these approach, the Gray-Level Co-occurrence Matrix (GLCM) has been widely applied because it captures texture information by quantifying the spatial relationships between pixel intensities. For example, a recent study proposed a fuzzy logic-based classification model that leverages GLCM features combined with particle swarm optimization to effectively categorize musculoskeletal radiographs (MURA) as normal or abnormal, achieving promising results with reduced computational resources [31]. Similarly, another study applied texture analysis using Haralick's and GLCM-based features on 3D brain MRI scans, demonstrating that measures such as contrast, correlation, energy, homogeneity, and entropy are highly useful for early tumor detection [32]. Further evidence comes from breast cancer studies, where GLCM and Gray Level Run Length Matrix (GLRLM) features extracted from fine needle biopsy microscopic images were classified using a KNN classifier, achieving up to 90% accuracy in distinguishing benign from malignant tumors [33].

The strength of GLCM lies in its ability to capture the spatial arrangement of pixel intensities, providing descriptive statistics such as contrast, correlation, and homogeneity that reflect the texture of an image, as originally formalized by Haralick and colleagues in 1973 [34]. These findings across different imaging

modalities highlight the effectiveness of GLCM features in extracting meaningful patterns, thereby justifying their role as a strong baseline before transitioning to deep learning methods such as convolutional neural networks (CNNs). In addition, a hybrid method combining Principal Component Analysis (PCA) with GLCM for feature extraction, followed by classification using a Probabilistic Neural Network-Radial Basis Function (PNN-RBF) model, has been proposed for brain tumor MRI images, achieving effective discrimination between normal, benign, and malignant cases [35]. More specifically, GLCM characterizes image texture by computing how often pairs of pixels with specific gray-level values occur in a defined spatial relationship, forming a matrix from which statistical measures such as contrast, correlation, and homogeneity can be derived [18]. Recent research on colorectal cancer further demonstrated that 3D GLCM features extracted from multiple color space channels, when combined with machine learning classifiers such as QDA and SVM, achieved detection accuracies above 97%, confirming the continued strength of handcrafted texture descriptors in tumor analysis [36].

While these handcrafted approaches are powerful and interpretable, their reliance on predefined statistical measures limits their adaptability to complex data distributions. By contrast, deep learning models such as CNNs automatically learn hierarchical representations from raw images, often surpassing handcrafted methods in large-scale datasets. Thus, GLCM features serve as an effective and interpretable foundation, while CNN-based approaches provide scalability and improved performance when sufficient data and computational resources are available.

III. METHODOLOGY

A. Data Description

This study used a dataset from Mendeley Data that consists of 296 multidetector computed tomography (MDCT) scans of the foot and ankle, encompassing both healthy and tumor-affected cases. The dataset is divided into two categories, with 134 images labeled as normal (healthy) and 162 images labeled as abnormal (tumor-present). Each image is stored in JPEG (.jpg) format with a standardized resolution of 512 × 512 pixels, ensuring consistency across all samples. The images were originally obtained in MDCT format and subsequently converted to JPEG for ease of handling and computational processing.

All MDCT scans in this study were initially read using the OpenCV library. Since the images were originally in RGB format, they were converted into grayscale prior to feature extraction to ensure compatibility with the GLCM method. The grayscale conversion represents each pixel intensity on a scale from 0 to 255, thereby preserving structural information while reducing computational complexity. This preprocessing step ensured that texture features could be extracted consistently across the entire dataset.

B. Feature Extraction (GLCM)

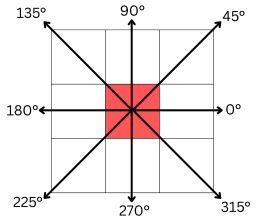

Fig. 1. GLCM orientations

In this study, feature extraction was performed using the Gray Level Co-occurrence Matrix (GLCM), a statistical method widely applied in medical imaging for texture analysis. GLCM captures spatial relationships between pixel intensities, thereby enabling the quantification of texture characteristics that distinguish healthy tissue from tumor-affected regions in MDCT scans. From the computed matrices, ten descriptors were extracted to provide a comprehensive representation of image texture: contrast (local intensity variation), dissimilarity (differences between neighboring gray levels), homogeneity (smoothness of texture), angular second moment (ASM) and energy (measures of textural uniformity), correlation (linear dependency of gray levels), mean, variance, standard deviation, and entropy (randomness or complexity of pixel distribution). These features, collectively, serve as discriminative indicators of underlying tissue properties and constitute the input for the classification models developed in this study. In particular, contrast, homogeneity, dissimilarity, energy, and correlation serve as the primary texture descriptors, while additional features such as mean, variance, standard deviation, and entropy were engineered from the co-occurrence matrix to enrich the feature space and enhance the discriminative capacity of the models.

To ensure a robust representation of spatial texture, GLCM matrices were computed across multiple orientations, specifically at $0°$, $45°$, $90°$, and $135°$, as illustrated in Fig. 1. This directional variation enables the extraction process to capture anisotropic textural patterns, thereby improving the discriminative capacity of the features for tumor detection tasks. Each feature relies on the co-occurrence probability $p(i, j, d, \theta)$, which represents the likelihood that a gray level i at a reference pixel occurs with gray level j at a neighbor pixel separated by distance d and orientation θ. Here, i and j denote gray-level intensities, d is the pixel distance, and θ represents the orientation angle [49].

Contrast measures the intensity difference between neighboring pixels, highlighting texture variations.

$$\text{Contrast} = \sum_i \sum_j (i-j)^2 p(i, j, d, \theta) \qquad (1)$$

While Dissimilarity quantifies the degree of local variation, similar to contrast, but using the absolute difference.

$$\text{Dissimilarity} = \sum_i \sum_j p(i,j,d,\theta)|i-j| \quad (2)$$

Homogeneity evaluates the closeness of the distribution to the GLCM diagonal, assigning higher weights to near-diagonal elements.

$$\text{Homogeneity} = \sum_i \sum_j \frac{p(i,j,d,\theta)}{1+|i-j|} \quad (3)$$

Textural uniformity reflected energy, which is calculated as the squared sum of the GLCM probabilities.

$$\text{Energy} = \sum_i \sum_j p(i,j,d,\theta)^2 \quad (4)$$

ASM is related to energy and measures image smoothness.

$$\text{ASM} = \sqrt{\text{Energy}} \quad (5)$$

Image complexity or randomness is captured by Entropy, with higher values indicating more disorder.

$$\text{Entropy} = -\sum_i \sum_j p(i,j,d,\theta)\ \ln\big(p(i,j,d,\theta)\big) \quad (6)$$

Finally, correlation assesses the linear dependency between gray levels of neighboring pixels.

$$\text{Correlation} = \sum_i \sum_j \left(ij\, p(i,j,d,\theta) - \mu_x \mu_y\right) \quad (7)$$

where μ_x, μ_y are the mean gray levels and σ_x, σ_y are the corresponding standard deviations.

In addition to these, statistical descriptors were also extracted. The mean represents the average gray-level value weighted by its co-occurrence probability.

$$\text{Mean} = \sum_i i * p(i,j,d,\theta) \quad (8)$$

Variance quantifies the spread of gray levels around the mean, providing a measure of texture heterogeneity.

$$\text{Variance} = \sum_i p(i,j,d,\theta)(i-mean)^2 \quad (9)$$

Finally, the standard deviation is defined as the square root of the variance, serving as a complementary measure of dispersion in the texture distribution.

$$\text{std} = \sqrt{variance} \quad (10)$$

C. Feature Scaling

In this research, feature scaling was applied to standardize the extracted GLCM features before they were used for classification. Specifically, z-score normalization was employed, which transforms each feature by subtracting its mean and dividing by its standard deviation. Z-score normalization (also known as Standardization) transforms data to have a mean of 0 and a unit variance [37], [38].

$$X_{\text{norm}} = \frac{X-\mu}{\sigma} \quad (11)$$

This process ensures that all features have a mean of zero and a standard deviation of one, thereby reducing the influence of varying scales among different feature metrics. Normalization not only facilitates a fair comparison across features but also improves the stability and convergence of machine learning models, particularly those sensitive to feature magnitude differences. By applying z-score normalization, the extracted texture features were placed on a common scale, enhancing their effectiveness in discriminating between healthy and tumor-affected foot MDCT scans.

D. Classification Model

To evaluate the effectiveness of the extracted GLCM features, we employed a set of widely used machine learning algorithms, representing different paradigms of classification. Linear models were represented by Logistic Regression, which models the probability of class membership based on a linear combination of input features and is well-suited for binary classification tasks [39]. Tree-based models included Decision Tree, which recursively partitions the data space based on feature thresholds [40], and Random Forest, an ensemble method that aggregates multiple decision trees to improve generalization and mitigate overfitting [41]. Additionally, Extreme Gradient Boosting (XGBoost) was utilized as an advanced gradient boosting framework, known for its efficiency and ability to handle complex, structured data [42]. From the probabilistic family, we applied Naïve Bayes, which relies on Bayes' theorem under the assumption of feature independence, offering computational simplicity and robustness in certain scenarios [43]. A distance-based approach was represented by k-Nearest Neighbors (KNN), which classifies instances based on the majority class of their nearest neighbors in the feature space [44]. Finally, we included the Support Vector Machine (SVM), a margin-based classifier that seeks an optimal hyperplane to maximize class separation, particularly effective in high-dimensional feature spaces [45].

By employing this diverse set of models, we aimed to conduct a comprehensive comparative analysis, allowing us to assess the relative strengths of linear versus non-linear, probabilistic versus distance-based, and

single versus ensemble approaches for the detection of foot tumors from MDCT scans.

E. Evaluation Metrics

The performance of the classification models was assessed using multiple evaluation metrics to capture different aspects of predictive capability. First, we utilized the confusion matrix, which provides a detailed breakdown of true positives, true negatives, false positives, and false negatives, thereby offering insights into classification accuracy and error distribution [46].

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{12}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{13}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{14}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{15}$$

$$\text{TPR} = \text{Recall} \tag{16}$$

$$\text{FPR} = (1 - \text{TPR}) \tag{17}$$

To further evaluate discriminative ability, we calculated the Area Under the Receiver Operating Characteristic Curve (ROC AUC), which reflects the trade-off between sensitivity and specificity across different thresholds and is widely regarded as a robust measure for binary classification tasks [47].

$$\text{AUC} = \sum_{k=1}^{n-1} \frac{(FPR_{k+1} - FPR_k) \cdot (TPR_k + TPR_{k+1})}{2} \tag{18}$$

In addition, the coefficient of determination (R-squared) was used to quantify the proportion of variance in the dependent variable explained by the model, offering an additional perspective on model fit and predictive consistency [48].

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (X_i - Y_i)^2}{\sum_{i=1}^{m} (\overline{Y} - Y_i)^2} \tag{19}$$

In this equation, X_i denotes the predicted value for the i-th observation, Y_i represents the actual ground-truth value, and $\overline{Y}$ is the mean of the observed values. The numerator, $\sum_{i=1}^{m} (X_i - Y_i)^2$, represents the residual sum of squares (RSS), which measures the discrepancy between the predictions and the true values. The denominator, $\sum_{i=1}^{m} (\overline{Y} - Y_i)^2$, corresponds to the total sum of squares (TSS), indicating the total variance present in the data. Thus, R^2 ranges from 0 to 1, where values closer to 1 indicate that a larger proportion of the variance in the dependent variable is accounted for by the model, while values near 0 indicate poor explanatory power [50].

To ensure reliability and mitigate overfitting, model evaluation was conducted using stratified a 5-fold cross-validation strategy, whereby the dataset was partitioned into five equal subsets. Each model was iteratively trained on four subsets and validated on the remaining one, with the process repeated five times to cover all folds. The final performance for each metric was reported as the average value across the five folds, thereby providing a more generalized and stable estimate of model performance.

IV. RESULT

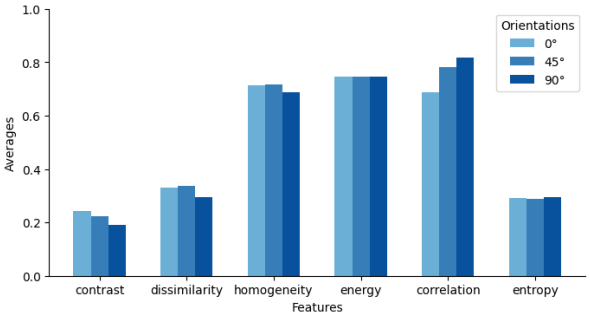

Fig. 2. Individual feature average per orientation

We extracted the GLCM features on three angles: 0, 45, and 90 degrees at a distance of one, giving rise to features with three different orientations each. The features stemmed from 296 data are investigated to procure holistic comprehensive insight. To this end, we scaled the features under MinMax scaler function and gathered the feature orientation average individually. Fig. 2 highlights a notable disparity in average per orientation among the features studied, revealing that contrast, dissimilarity, homogeneity, and correlation are more varied than energy and entropy.

Table 1. Average correlation of features against targets

GLCM features	Correlation coefficients
Std	0.88
Mean	0.80
Variance	0.80
Entropy	0.79
Dissimilarity	0.73
Contrast	0.69
Correlation	0.63
Homogeneity	-0.78
Energy	-0.79
ASM	-0.81

We explored further the correlation between the features and targets by averaging the correlation coefficient score deduced under three angles. Table 1 itemized a detailed tabulation of the scores within the range -1 to 1 systematically. According to the table, metrics such as homogeneity, energy, and ASM have negative correlations, whereas the remaining exhibit positive correlations towards targets.

All models advanced in this research originated from state-of-the-art sklearn. We trained the determined models under scaled features, demonstrating the result in Table 2. The table suggests evaluation scores across defined metrics. We computed models

Table 2. Models performance

Models	Acc	Prec	Rec	F1	AUC	r^2
Logistic Regression	1.0	1.0	1.0	1.0	1.0	1.0
Decision Tree	0.99	0.99	0.99	0.99	0.99	0.97
Random Forest	1.0	1.0	1.0	1.0	1.0	1.0

average performance across folds since we managed the training in a five-fold cross-validation fashion. As the table records, all models achieved significant performance from 0.97 to 1.0 regarding the metrics.

V. Discussions

We captured the GLCM features at 0°, 45°, and 90° to preserve the features unique, as initial assessments demonstrated that the remaining angle within the range of 135° to 315° by 45° in difference yields overlap features towards those three. In order to enhance the interpretability of Fig. 1, we bifurcate the features based on their average scores. Features exhibiting a score less than 0.5 are posited to the low-average category, whereas those with scores exceeding or equal to 0.5 are posited to the high-average category. The energy at any orientations shows scores resembling 0.75, suggesting high uniformity on co-occurrence pixels at an offset of 1. The presence of repetitive patterns, in which numerous similar neighbor pairs appear with similar frequencies, leads to a diminished entropy score of approximately 0.29 at any orientations. The number implies low information complexity. The high correlation averages at 90° orientation provide notable evidence for the strong dependency between perpendicular pixels. Simultaneously, the extracted features at 0° and 45° orientation are lower yet relatively significant compared to others. The contrast and correlation have a reverse relationship due to the distribution of gray co-occurrence matrix elements that fall close to the diagonal. It indicates that most data have smooth surfaces signified by marginal grey intensity. Furthermore, the negative association of dissimilarity versus homogeneity also corroborates the findings.

The aforementioned features are the GLCM key features. However, given the relatively small dataset, we engineered a few features originating from gray co-occurrence matrix elements and GLCM features to expand the feature dimensions, yielding improved generalizability. The feature engineering produced additional features, which are present in Table 1 but absent in Fig. 1. Table 1 reveals how individual features bond to the target. The negative relationship nature of homogeneity, energy, and ASM indicates that the lower the value of the features, the more the data tends to be recognized as a healthy foot. Conversely, the data with the remaining features are distinguished as metastasized ones. Nonetheless, it is crucial to note that their association does not establish causality for data to hold a specific target.

The feature engineering brought to boost sample efficiency has proven to improve model performance. In addition, parameter tuning is also instrumental in achieving those performances. We did not find our model experiencing overfitting based on the parameter tuning experiment. For validation purposes, we subjected the per-fold performance of models, preceding computing the performance in average. Individual model performance shows that a significant gap in per-fold performance did not emerge but merely 0.02 maximum for confusion matrix, AUC, and R-squared. The profound performance implicitly stated that the computation behaviour of the specified models learns the data effectively. The notable performance of a logistic regression, illustrates that the sigmoid decision boundaries by threshold of 0.5 operated reliably. Ultimately, the 296x30 feature size in this study is considered small-to-medium so that data non-linearity is effectively handled.

VI. Conclusions

We conclude that all results and findings are worth considering as the main reference on tumor detection application advancement, especially foot tumors. Although the latest tumor detection study ubiquitously addresses the detection under a combined healthy and tumor variant image dataset, our research paves an alternative way for two-stage foot tumor detection. In this research, we propose the first stage of the method. This research procures compelling evidence that the specified models are adept at resolving presence-absence foot tumor detection. Moreover, our study underscores how the applied feature engineering on three orientations and designated offset did not hinder the determined models.

Despite the excellence, we encourage further research to harness filter-based methods such as Convolutional Neural Networks (CNN) to excavate the knowledge in the GLCM versus CNN comparison. Besides, RNN-based models embody temporal dependencies, and attention-based transformer models remain underexplored for such a detection task. Nevertheless, the mentioned models might require a larger amount of data to attain top-notch performance, so data augmentation or labeled data gathering is indispensable.

Acknowledgment

I sincerely thank Telkom University for its generous support throughout this research.

References

[1] V. A. Singh, V. Sandhu, C. T. Yong, and N. F. Yasin, "Tumours of the foot: A 10 years retrospective analysis," J. Orthop. Surg., vol. 32, no. 1, pp. 10225536241248706, 2024.

[2] S. Rammelt, H. Fritzsche, C. Hofbauer, and others, "Malignant tumours of the foot and ankle," Foot Ankle Surg., vol. 26, pp. 363–370, 2020.

[3] A. Bhangu, J. A. S. Beard, and R. J. Grimer, "Should soft tissue sarcomas be treated at a specialist centre?," Sarcoma, vol. 8, pp. 1–6, 2004.

[4] H. Zwipp and S. Rammelt, "Tumoren des Fusses (foot tumours)," in Tscherne Unfallchirurgie Fuss, H. Zwipp and S. Rammelt, Eds. Heidelberg, Germany: Springer, 2014, pp. 755–770.

[5] H. DeGroot III, "Approach to the management of soft tissue tumors of the foot and ankle," Foot Ankle Spec., vol. 1, no. 3, pp. 168–176, 2008.

[6] A. K. Walling and S. I. Gasser, "Soft-tissue and bone tumors about the foot and ankle," Clin. Sports Med., vol. 13, no. 4, pp. 909–938, 1994.

[7] M. O. Karaca, K. Başarır, A. Merter, E. Acar, E. A. Özbek, M. Özyıldıran, and H. Y. Yıldız, "Malignant tumors of the foot and ankle," Foot Ankle Int., vol. 43, no. 9, pp. 1232–1241, 2022.

[8] K. Suzuki, "Overview of deep learning in medical imaging," Radiol. Phys. Technol., vol. 10, pp. 257–273, 2017.

[9] L. Šajn and M. Kukar, "Image processing and machine learning for fully automated probabilistic evaluation of medical images," Comput. Methods Programs Biomed., vol. 104, no. 3, pp. e75–e86, 2011.

[10] M. Nazir, S. Shakil, and K. Khurshid, "Role of deep learning in brain tumor detection and classification (2015 to 2020): A review," Comput. Med. Imag. Graph., vol. 91, p. 101940, 2021.

[11] S. Gull and S. Akbar, "Artificial intelligence in brain tumor detection through MRI scans: Advancements and challenges," Artif. Intell. Internet Things, pp. 241–276, 2021.

[12] S. Hossain, A. Chakrabarty, T. R. Gadekallu, M. Alazab, and M. J. Piran, "Vision transformers, ensemble model, and transfer learning leveraging explainable AI for brain tumor detection and classification," IEEE J. Biomed. Health Inform., vol. 28, no. 3, pp. 1261–1272, 2023.

[13] G. Kasinathan and S. Jayakumar, "Cloud-based lung tumor detection and stage classification using deep learning techniques," Biomed. Res. Int., vol. 2022, no. 1, p. 4185835, 2022.

[14] R. Y. Lee, Y. Wu, D. Goh, V. Tan, C. W. Ng, J. C. T. Lim, M. C. Lau, and J. P. S. Yeong, "Application of artificial intelligence to in vitro tumor modeling and characterization of the tumor microenvironment," Adv. Healthcare Mater., vol. 12, no. 14, p. 2202457, 2023.

[15] A. S. Vironicka and J. G. R. Sathiaseelan, "Texture feature extraction with medical image using GLCM and machine learning techniques," ICTACT J. Image Video Process., vol. 12, no. 4, 2022.

[16] M. Benco, R. Hudec, P. Kamencay, M. Zachariasova, and S. Matuska, "An advanced approach to extraction of colour texture features based on GLCM," Int. J. Adv. Robot. Syst., vol. 11, no. 7, p. 104, 2014.

[17] M. R. I. Hossain, I. Ahmed, and M. H. Kabir, "Automatic lung tumor detection based on GLCM features," in Asian Conf. Comput. Vision, 2014, pp. 109–121.

[18] N. Zulpe and V. Pawar, "GLCM textural features for brain tumor classification," Int. J. Comput. Sci. Issues, vol. 9, no. 3, p. 354, 2012.

[19] G. Dheepak, D. Vaishali, and others, "Brain tumor classification: A novel approach integrating GLCM, LBP and composite features," Front. Oncol., vol. 13, p. 1248452, 2024.

[20] A. Toepfer, N. Harrasser, M. Recker, and others, "Distribution patterns of foot and ankle tumors: A university tumor institute experience," BMC Cancer, vol. 18, p. 735, 2018.

[21] A. Toepfer, "Tumors of the foot and ankle: A review of the principles of diagnostics and treatment," Fuss Sprunggelenk, vol. 15, no. 2, pp. 82–96, 2017.

[22] L. H. Wetzel and E. Levine, "Soft-tissue tumors of the foot: Value of MR imaging for specific diagnosis," AJR Amer. J. Roentgenol., vol. 155, no. 5, pp. 1025–1030, 1990.

[23] S. Waldt, H. Rechl, E. J. Rummeny, and others, "Imaging of benign and malignant soft tissue masses of the foot," Eur. Radiol., vol. 13, pp. 1125–1136, 2003.

[24] A. Mohseni, E. Ghotbi, F. Kazemi, A. Shababi, S. C. Jahan, A. Mohseni, and N. Shababi, "Artificial intelligence in radiology: What is its true role at present, and where is the evidence?," Radiol. Clin., vol. 62, no. 6, pp. 935–947, 2024.

[25] J. M. Boone, "Radiological interpretation 2020: Toward quantitative image assessment," Med. Phys., vol. 34, no. 11, pp. 4173–4179, 2007.

[26] W. Chen, Y. Zhang, X. Wang, and J. Liu, "Comparative analysis of imaging modalities for diagnosing musculoskeletal disorders," J. Innov. Med. Res., vol. 3, no. 1, pp. 45–53, 2024.

[27] D. S. Sule and R. Ampofo, "Diagnostic imaging techniques: Current trends and challenges," Diagn. Adv. Precis. Med. Drug Develop., pp. 49–66, 2025.

[28] D. Ballard and J. Sklansky, "Tumor detection in radiographs," Comput. Biomed. Res., vol. 6, no. 4, pp. 299–321, 1973.

[29] H.-X. Zhao, C.-X. Xia, H.-X. Yin, N. Guo, and Q. Zhu, "The value and limitations of contrast-enhanced transrectal ultrasonography for the detection of prostate cancer," Eur. J. Radiol., vol. 82, no. 11, pp. e641–e647, 2013.

[30] B. C. Jones and L. M. Fayad, "Musculoskeletal tumor imaging: Focus on emerging techniques," Sem. Roentgenol., vol. 52, no. 4, pp. 269–281, 2017.

[31] V. Narayan, P. K. Mall, S. Awasthi, S. Srivastava, and A. Gupta, "FuzzyNet: Medical image classification based on GLCM texture feature," in Proc. Int. Conf. Artif. Intell. Smart Commun. (AISC), 2023, pp. 769–773.

[32] B. Dhruv, N. Mittal, and M. Modi, "Study of Haralick's and GLCM texture analysis on 3D medical images," Int. J. Neurosci., vol. 129, no. 4, pp. 350–362, 2019.

[33] P. Filipczuk, T. Fevens, A. Krzyżak, and A. Obuchowicz, "GLCM and GLRLM based texture features for computer-aided breast cancer diagnosis," J. Med. Inform. Technol., vol. 19, 2012.

[34] M. Hall-Beyer, "GLCM texture: A tutorial v. 3.0 March 2017," 2017.

[35] S. S. Shirke, J. A. Kendule, and S. G. Vyawhare, "An approach for PCA and GLCM based MRI image classification," in Techno-Societal 2016: Int. Conf. Adv. Technol. Soc. Appl., 2016, pp. 265–274.

[36] A. M. Alqudah and A. Alqudah, "Improving machine learning recognition of colorectal cancer using 3D GLCM applied to different color spaces," Multimed. Tools Appl., vol. 81, pp. 10839–10860, 2022.

[37] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Waltham, MA: Elsevier, 2011. [Online]. Available: https://shop.elsevier.com/books/data-mining-concepts-and-techniques/han/978-0-12-381479-1

[38] A. Reverter, W. Barris, S. McWilliam, K. A. Byrne, Y. H. Wang, S. H. Tan, N. Hudson, and B. P. Dalrymple, "Validation of alternative methods of data normalization in gene co-expression studies," Bioinformatics, vol. 21, no. 7, pp. 1112–1120, 2004.

[39] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression, 3rd ed. Hoboken, NJ: Wiley, 2013.

[40] J. R. Quinlan, "Induction of decision trees," Mach. Learn., vol. 1, no. 1, pp. 81–106, 1986.

[41] L. Breiman, "Random forests," Machine Learn., vol. 45, no. 1, pp. 5–32, 2001.

[42] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2016, pp. 785–794.

[43] I. Rish, "An empirical study of the naive Bayes classifier," in Proc. IJCAI Workshop Empirical Methods AI, 2001.

[44] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theory, vol. 13, no. 1, pp. 21–27, 1967.

[45] C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, 1995.

[46] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," J. Mach. Learn. Technol., vol. 2, no. 1, pp. 37–63, 2011.

[47] T. Fawcett, "An introduction to ROC analysis," Pattern Recognit. Lett., vol. 27, no. 8, pp. 861–874, 2006.

[48] N. J. Nagelkerke, "A note on a general definition of the coefficient of determination," Biometrika, vol. 78, no. 3, pp. 691–692, 1991.

[49] F. Utaminingrum, A. M. Alqadri, I. K. Somawirata, C. Karim, A. Septiarini, C.-Y. Lin, and T. K. Shih, "Feature selection of gray-level Cooccurrence matrix using genetic algorithm with Extreme learning machine classification for early detection of Pole roads," Results in Engineering, vol. 20, p. 101437, 2023.

[50] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," PeerJ Comput. Sci., vol. 7, p. e623, 2021.