A COMPUTER-AIDED DIAGNOSIS SYSTEM FOR KNEE OSTEOARTHRITIS USING DEEP LEARNING AND X-RAY IMAGES

JEEVITHA RAVULA[1], M. SAHITHI[2], CH. RAGHAVENDRA[3], PRABHAKAR CHAUBEY[4], P. SHIVA MANI[5]

ASSISTANT PROFESSOR[1], UG SCHOLAR[2,3,4&5]

DEPARTMENT OF AI&ML, NARSIMHA REDDY ENGINEERING COLLEGE (UGC- AUTONOMOUS) MAISAMMAGUDA (V), KOMPALLY, SECUNDERABAD, TELANGANA-500100

ABSTRACT

Knee osteoarthritis (OA) is a prevalent degenerative joint disease that significantly impacts the quality of life. Early and accurate diagnosis is critical for effective management and treatment. This project, titled "Automated Knee Osteoarthritis Prediction and Classification from X-ray Images Using Deep Learning," presents a novel approach for diagnosing and grading knee OA using deep learning models. The system is developed using Python for backend computation, Flask as the web framework, and HTML, CSS, and JavaScript for the frontend. Two state-of-the-art deep learning models, VGG16 and MobileNetV2, have been employed to classify X-ray images of knee joints into five categories based on the Kellgren and Lawrence grading system: Normal, Doubtful, Mild, Moderate, and Severe. The dataset consists of 1,650 high-quality 8-bit grayscale X-ray images collected from reputable hospitals and diagnostic centers using the PROTEC PRS 500E X-ray machine. Each image has been manually annotated by two medical experts for accuracy and reliability. The VGG16 model achieved an impressive train accuracy of 93% and a test accuracy of 92%, while the MobileNetV2 model outperformed with a train accuracy of 96% and a test accuracy of 96%. Performance metrics, including Accuracy, Precision, Recall, F-Measure, and a Confusion Matrix, have been computed to evaluate each model comprehensively. The findings of this study highlight the potential of deep learning in automating the detection and grading of knee osteoarthritis, offering a cost-effective and scalable solution for clinical use. The system's superior performance, particularly with the MobileNetV2 model, demonstrates its applicability in real-world diagnostic settings.

INTRODUCTION

Knee osteoarthritis (KOA) is a chronic degenerative joint disorder that predominantly affects the knee, one of the body's most crucial weight-bearing joints. It is a leading cause of disability worldwide, especially among middle-aged and elderly populations. KOA occurs due to the gradual breakdown of cartilage, the protective tissue covering the ends of bones in the knee joint, leading to pain, stiffness, swelling, and reduced mobility. The progression of KOA is influenced by various factors, including aging, obesity, genetic predisposition, joint injuries, and repetitive stress on the knees from physical activities. The condition is typically classified using the Kellgren and Lawrence grading system, which divides KOA into five grades: Normal, Doubtful, Mild, Moderate, and Severe, based on X-ray findings such as joint space narrowing, osteophyte formation, and bone deformities. Early diagnosis and treatment of KOA are crucial to managing symptoms, preventing further joint damage, and improving patients' quality of life. However, traditional diagnostic methods, which rely on manual evaluation of X-ray images by radiologists, can be time-consuming and subjective. This has led to a growing interest in automated approaches, such as deep learning, to enhance diagnostic accuracy, reduce variability, and support clinical decision-making. With the increasing prevalence of knee osteoarthritis due to aging populations and lifestyle changes, the development of efficient and reliable diagnostic systems is more critical than ever. Such advancements hold the potential to revolutionize the management of KOA, enabling timely interventions and better long-term outcomes for patients.

LITERATURE SURVEY

1) Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies

AUTHORS: A. Cui, H. Li, D. Wang, J. Zhong, Y. Chen, and H. Lu,

Background: Knee osteoarthritis (OA) is a major cause of disability in the elderly, however, there are few studies to estimate the global prevalence, incidence, and risk factors of knee OA.

Methods: For this study, we searched PUBMED, EMBASE and SCOPUS from inception to April 4, 2020, without language restriction. We identified eligible studies with information on the prevalence or incidence of knee OA in population-based observational studies and extracted data from published reports. We did random-effects meta-analysis to generate estimates. This study was registered with PROSPERO (CRD42020181035).

Findings: Out of 9570 records identified, 88 studies with 10,081,952 participants were eligible for this study. The pooled global prevalence of knee OA was 16·0% (95% CI, 14·3%-17·8%) in individuals aged 15 and over and was 22·9% (95% CI, 19·8%-26·1%) in individuals aged 40 and over. Correspondingly, there are around 654·1 (95% CI, 565·6-745·6) million individuals (40 years and older) with knee OA in 2020 worldwide. The pooled global incidence of knee OA was 203 per 10,000 person-years (95% CI, 106-331) in individuals aged 20 and over.

Correspondingly, there are around annual 86·7 (95% CI, 45·3-141·3) million individuals (20 years and older) with incident knee OA in 2020 worldwide. The prevalence and incidence varied substantially between individual countries and increased with age. The ratios of prevalence and incidence in females and males were 1·69 (95% CI, 1·59-1·80, p<0·00) and 1·39 (95% CI, 1·24-1·56, p<0·00), respectively.

Interpretation: Our study provides the global prevalence (16·0% [95% CI, 14·3%-17·8%]) and incidence (203 per 10,000 person-years [95% CI, 106-331]) of knee OA. These findings can be used to better assess the global health burden of knee OA. Further prospective cohort studies are warranted to identify modifiable risk factors for providing effectively preventive strategies in the early stages of the disease.

2) A health-impact assessment of an ergonomic measure to reduce the risk of work-related lower back pain, lumbosacral radicular syndrome and knee osteoarthritis among floor layers in The Netherlands

AUTHORS: P. P. F. M. Kuijer, H. F. van der Molen, and S. Visser

Sand–cement-bound screed floor layers are at risk of work-related lower back pain, lumbosacral radicular syndrome and knee osteoarthritis, given their working technique of levelling screed with their trunk bent while mainly supported by their hands and knees. To reduce the physical demands of bending of the trunk and kneeling, a manually movable screed-levelling machine was developed for floor layers in the Netherlands. The aim of this paper is to estimate the potential health gains of working with a manually movable screed-levelling machine on the risk of lower back pain (LBP), lumbosacral radicular syndrome (LRS) and knee osteoarthritis (KOA) compared to traditional working techniques. This potential health gain was assessed using the epidemiological population estimates of the Population Attributable Fraction (PAF) and the Potential Impact Fraction (PIF), combined with work-related risk estimates for these three disorders from systematic reviews. The percentage of workers exceeding these risk estimates was based on worksite observations among 28 floor layers. For LBP, 16/18 workers were at risk when using traditional working techniques, with a PAF = 38%, and for those using a manually movable screed-levelling machine, this was 6/10 with a PIF = 13%. For LRS, these data were 16/18 with a PAF = 55% and 14/18 with a PIF = 18%, and for KOA, 8/10 with a PAF = 35% and 2/10 with a PIF = 26%. A manually movable screed-levelling machine might have a significant impact on the prevention of LBP, LRS and KOA among floor layers in the Netherlands, and health-impact assessments are a feasible approach for assessing health gains in an efficient way.

3) Evaluating the efficacy of deep learning models for knee osteoarthritis prediction based on kellgrenlawrence grading system

AUTHORS: V. K. V, V. Kalpana, and G. H. Kumar

Osteoarthritis of the knee, also known as OA has been determined that osteoarthritis of the knee is the leading cause of activity limitations and the development of disability, particularly in people who are older. The utilisation of artificial intelligence (AI) methodologies grounded in deep learning (DL) has yielded promising outcomes in the realm of radiographic interpretation. The utilisation of deep learning in the healthcare industry has yielded remarkable outcomes and elevated the benchmark for the quality of medical treatment. This study used knee OA as a clinical scenario to compare twelve transfer learning DL models for detecting the grade of KOA from a radiograph, compared their accuracy, and determined the best model for detecting KOA. The models exhibited a range of 30% to 98% in detecting the KOA. It was determined that MobileNet was responsible for the highest level of accuracy, which came in at 98.36%. It has high training and validation accuracy. The maximum loss was observed for EfficientNetB7. DL approaches created by skilled radiologists and orthopaedic specialists could help smaller hospitals learn and make more emergency room. This would be especially helpful in situations when medical personnel may not be available.

4) Prediction of the Rapid Progression of Knee Osteoarthritis Using Automated Machine Learning: A Novel Precision Health Approach for Chronic Degenerative Disease

AUTHORS: S. Castagno, M. Birch, M. van der Schaar, and A. McCaskie

Objectives: To facilitate the stratification of patients with osteoarthritis (OA) for new treatment development and clinical trial recruitment, we created an automated machine learning (autoML) tool predicting the rapid progression of knee OA over a 2-year period.

Methods: We developed autoML models integrating clinical, biochemical, X-ray and MRI data. Using two data sets within the OA Initiative-the Foundation for the National Institutes of Health OA Biomarker Consortium for training and hold-out validation, and the Pivotal Osteoarthritis Initiative MRI Analyses study for external validation-we employed two distinct definitions of clinical outcomes: Multiclass (categorising OA progression into pain and/or radiographic) and binary. Key predictors of progression were identified through advanced interpretability techniques, and subgroup analyses were conducted by age, sex and ethnicity with a focus on early-stage disease.

Results: Although the most reliable models incorporated all available features, simpler models including only clinical variables achieved robust external validation performance, with area under the precision-recall curve (AUC-PRC) 0.727 (95% CI: 0.726 to 0.728) for multiclass predictions; and AUC-PRC 0.764 (95% CI: 0.762 to 0.766) for binary predictions. Multiclass models performed best in patients with early-stage OA (AUC-PRC 0.724-0.806) whereas binary models were more reliable in patients younger than 60 (AUC-PRC 0.617-0.693). Patient-

reported outcomes and MRI features emerged as key predictors of progression, though subgroup differences were noted. Finally, we developed web-based applications to visualise our personalised predictions.

Conclusions: Our novel tool's transparency and reliability in predicting rapid knee OA progression distinguish it from conventional 'black-box' methods and are more likely to facilitate its acceptance by clinicians and patients, enabling effective implementation in clinical practice.

5) Machine learning based osteoarthritis detection methods in different imaging modalities: A review

AUTHORS: A. A. S. Afroze, R. Tamilselvi, and M. G. P. Beham Osteoarthritis (OA) is a bone disease that mainly affects the cartilage. Even though there are many diseases that are commonly noticed in bones, one of the most dangerous diseases is OA. The breakdown of the cartilage bone is the cause of OA. According to the survey given by the National Institute on Aging, it is revealed that most of the people in their old age are at the very advanced stage of OA. X-ray is the common imaging modality for analysing the severity of Osteoarthritis. When needed for advanced level of investigation, MRI scans and thermal images are also initialized. There are numerous methods for the analysis of OA from different modalities in the very early stage. These methods may be semi-automatic and automatic. But all the developed algorithms gave results based on the space width, and texture feature only and didn't provide any quantitative analysis based on any standard parameters. The main aim of this work is to present major research challenges in different OA detection methods, discuss different machine learning-based OA detection methods and analyse their performance. The research gap in the existing methods such as an empirical model for the detection of OA and the standard parameters for the measurement of bone marrow is discussed in the proposed paper.

SYSTEM ANALYSIS

EXISTING SYSTEM:

- The existing system for knee osteoarthritis (KOA) prediction and classification leveraged a highly advanced approach centered on a modified compact convolutional transformer model, KOA-CCTNet. This foundational model was designed to enhance feature extraction and classification capabilities, achieving a commendable test accuracy of 94.58%.
- To overcome the limitations of individual datasets, the system aggregated data from four different sources, generating an extensive dataset comprising 110,232 raw X-ray images. Data augmentation was employed using the deep convolutional generative adversarial network (DCGAN) to synthesize additional training samples, ensuring a diverse and balanced dataset.
- Advanced image preprocessing techniques were applied to improve image quality and optimize the model's performance. These included adaptive histogram equalization (AHE) to enhance contrast, fast non-local means to reduce noise, and image resizing for consistency. This meticulous preprocessing pipeline not only improved the clarity of the X-ray images but also ensured that the critical features required for accurate KOA classification were preserved.
- By combining KOA-CCTNet with a robust dataset and sophisticated preprocessing methods, the existing system laid a strong foundation for automating the diagnosis and grading of knee osteoarthritis.

DISADVANTAGES OF EXISTING SYSTEM:

- Complexity of Model Architecture: The modified compact convolutional transformer (KOA-CCTNet) employed in the existing system involves a highly complex architecture, which increases computational requirements and makes it challenging to deploy in resource-constrained environments such as small clinics or rural healthcare centers.
- Dependence on Large Datasets: The system relies on an extensive dataset of 110,232 images, which is generated by aggregating multiple datasets and employing DCGAN-based augmentation. This dependence on large-scale data increases the need for significant storage and computational resources, potentially limiting accessibility for smaller institutions.
- Time-Consuming Data Preprocessing: Advanced preprocessing techniques, such as adaptive histogram equalization (AHE) and fast non-local means, enhance image quality but require substantial processing time and expertise to implement effectively.
- Synthetic Data Limitations: While DCGAN-based augmentation generates diverse training samples, there is a potential risk that synthetic data may not fully capture the nuances of real-world X-ray images, which could impact the generalizability of the model.
- Scalability Concerns: The computationally intensive nature of KOA-CCTNet and the preprocessing pipeline may pose challenges when scaling the system for widespread use in healthcare applications or for real-time processing.
- Higher Resource Requirements: The complex nature of the model and preprocessing pipeline necessitates

high-performance hardware, which may not be readily available in all healthcare facilities, limiting its deployment in low-resource settings.

- These limitations highlight areas for improvement in achieving a more efficient, scalable, and universally applicable solution for knee osteoarthritis prediction and classification.

PROPOSED SYSTEM:

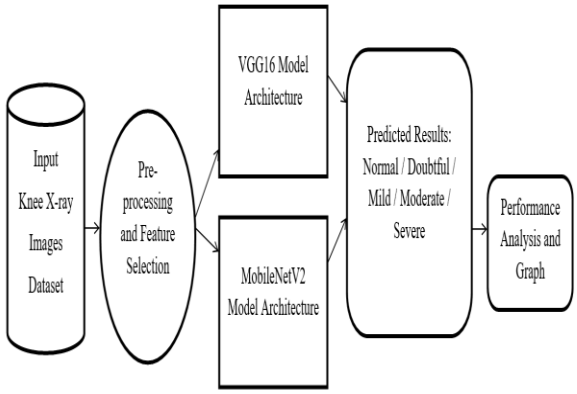
- The proposed system focuses on automating the prediction and classification of knee osteoarthritis (KOA) using deep learning techniques to provide a streamlined, accurate, and efficient solution for diagnostic purposes. It integrates cutting-edge machine learning algorithms with user-friendly web-based technology for seamless functionality.
- Development Framework:
  - Backend: Developed using Python, which provides extensive libraries for deep learning and image processing.
  - Web Framework: Flask is employed for creating a robust and lightweight web interface.
  - Frontend: HTML, CSS, and JavaScript are used for designing a responsive and intuitive user interface.
- Dataset Details: The system utilizes a dataset comprising 1,650 X-ray images of knee joints. These images were collected from reputable hospitals and diagnostic centers using the PROTEC PRS 500E X-ray machine.
  - Image Characteristics: The images are 8-bit grayscale X-rays.
  - Annotations: Each image is manually labeled by two medical experts based on the Kellgren and Lawrence grading system, ensuring reliability and consistency.
  - Classes: The dataset is divided into five categories: Normal, Doubtful, Mild, Moderate, and Severe.
- Region of Interest Extraction: A novel preprocessing method is implemented to automatically extract the cartilage region (region of interest) from the X-ray images. This technique is based on pixel density, isolating the critical area for analysis and enhancing model performance by focusing on relevant features.
- Deep Learning Models: The system integrates two advanced deep learning models:
  - VGG16 Model: Known for its simplicity and powerful feature extraction capabilities, this model achieved a training accuracy of 93% and a test accuracy of 92%.
  - MobileNetV2 Model: A lightweight yet efficient architecture optimized for mobile and embedded applications, achieving a superior training accuracy of 96% and a test accuracy of 96%.
- Performance Evaluation: Comprehensive performance metrics, including Accuracy, Precision, Recall, F-Measure, and a Confusion Matrix, are utilized to assess the efficiency of both models. These metrics ensure an in-depth understanding of the system's classification capabilities.
- Visualization and Analysis: The system incorporates graphical tools to visualize the performance metrics and results, allowing end-users to interpret and validate the system's predictions.
- By leveraging advanced deep learning techniques and a streamlined architecture, the proposed system aims to deliver an effective solution for KOA classification, ensuring clinical utility and scalability.

ADVANTAGES OF PROPOSED SYSTEM:

- Enhanced Prediction Accuracy: The use of advanced deep learning models, particularly MobileNetV2 and VGG16, ensures high prediction accuracy. With the MobileNetV2 model achieving 96% test accuracy, the system provides reliable classification of knee osteoarthritis into five distinct grades.
- Automated Region of Interest (ROI) Extraction: The proposed system incorporates a novel technique to automatically extract the cartilage region based on pixel density, streamlining the preprocessing pipeline and reducing manual intervention. This ensures that the analysis focuses on the most critical features of the X-ray images.
- Efficient Data Utilization: By using a high-quality dataset of 1,650 expertly annotated X-ray images, the system strikes a balance between dataset size and performance. The use of real-world, labeled data ensures the model's robustness without over-reliance on synthetic or augmented data.
- Comprehensive Evaluation Metrics: The inclusion of performance metrics such as Accuracy, Precision, Recall, F-Measure, and a Confusion Matrix provides a detailed understanding of model effectiveness, enabling transparent evaluation and validation.

- User-Friendly Web Interface: Built using Flask for the backend and HTML, CSS, and JavaScript for the frontend, the system provides a seamless and intuitive interface for healthcare professionals. The web-based deployment makes it accessible and easy to use in clinical settings.
- Scalability and Flexibility: The lightweight architecture of MobileNetV2 ensures the system can be deployed on a wide range of devices, including resource-constrained environments, without compromising performance.
- Improved Workflow Efficiency: The automation of KOA detection and classification reduces the burden on radiologists and clinicians, saving time and enabling them to focus on patient care.
- Visualized Results: Graphical representations of performance metrics and predictions enhance interpretability, making the system more effective for decision-making and reporting.
- Cost-Effective Solution: The system's ability to provide accurate KOA classification without the need for extensive hardware or infrastructure makes it a cost-efficient alternative for healthcare facilities.
- Real-World Applicability: The high performance and robust design of the system make it highly suitable for deployment in hospitals, diagnostic centers, and research institutions, supporting early detection and treatment of knee osteoarthritis.
- The proposed system effectively combines advanced technology and practical usability, ensuring it meets the needs of modern healthcare applications.

IMPLEMENTATION



MODULES:

Data Collection:

- In the first module of the Automated Knee Osteoarthritis Prediction and Classification from X-ray Images Using Deep Learning, we make the data collection process. This is the first real step towards the real development of a deep learning model, collecting data. This is a critical step that will cascade in how good the model will be, the more and better data that we get; the better our model will perform
- There are several techniques to collect the data, like web scraping, manual interventions. The dataset is located in the model folder. The dataset is referred from the popular dataset repository called kaggle. The following is the link of the dataset:
- Kaggle Dataset Link:

https://www.kaggle.com/datasets/jayaprakashpondy/knee-xray

Dataset:

- In this module, set up two main directories: training, and testing. Training, and testing directories, creates subdirectories for each class label. In this case, The five classes "Normal", "Doubtful", "Mild", "Moderate", "Severe"
- Place the images into their respective class directories based on their labels. Each image should be placed in the directory corresponding to its class label and by organizing the dataset in this structure, it becomes straightforward to load the data into your deep learning framework for training and testing purposes and Total dataset size is 2,645.

Data Preparation:

- During the data preparation stage, it is crucial to preprocess the data to ensure it is suitable for training. This involves tasks such as resizing images to a standard size, normalizing pixel values, and encoding labels if necessary. To achieve this, the ImageDataGenerator from Keras can be utilized. For instance, to resize images to a standard size of 224x224 pixels, the Target_size parameter can be set to (img_height, img_width) = (224, 224).
- Additionally, pixel values can be normalized by setting the rescale parameter to 1./255, which scales the pixel values to be between 0 and 1. Furthermore, data augmentation techniques such as random shear and zoom can be applied to enhance the training data. By leveraging these techniques, the data can be effectively preprocessed to improve the performance of the deep learning model.

Feature Extraction:

- For models like MobileNetV2 and VGG16, which come pre-trained with feature extraction layers, explicit feature extraction may not always be necessary. By setting trainable = False, we freeze these pre-trained layers, allowing them to retain their learned representations while preventing further updates during training.
- In the context of MobileNetV2 and VGG16, setting trainable = False ensures that the weights of the feature extraction layers remain fixed during training. This approach is commonly used in transfer learning scenarios, where the pre-trained model is fine-tuned on a new dataset for a specific task, such as image classification or object detection.
- By adopting this strategy, we strike a balance between leveraging powerful pre-trained representations and adapting the model to our specific dataset, ultimately improving both training efficiency and model performance.

Splitting the dataset:

- Divide your dataset into training and validation to evaluate your model's performance. Typically, you might use an 80-20 split, but this can vary based on your dataset size and specific requirements.

Model Selection:

- The training module is responsible for training the deep learning models using the preprocessed data. It implements two popular architectures: MobileNetV2 and VGG16.

MobileNetV2:

- MobileNetV2 is a lightweight and efficient convolutional neural network architecture designed for mobile and embedded devices. It uses depth-wise separable convolutions to build deep neural networks while maintaining a small model size and low computational complexity. MobileNetV2 consists of an initial fully convolutional layer, followed by a series of inverted residual blocks with linear bottlenecks. These blocks use depth-wise convolutions to filter features and 1x1 convolutions to combine features. The network ends with a final convolution layer and global average pooling.

VGG16:

- VGG16 is a widely used deep convolutional neural network architecture known for its straightforward design and strong performance on image classification tasks. Developed by the Visual Geometry Group (VGG) at Oxford, it consists of 16 layers with weights, specifically 13 convolutional layers and 3 fully connected layers. Each convolutional layer in VGG16 employs a small 3x3 kernel, stacked in multiple layers to progressively extract deeper and more complex features from input images. This uniform filter size simplifies the model while allowing it to effectively capture spatial hierarchies. VGG16 increases its depth by doubling the number of filters at each level, starting from 64 and reaching up to 512 filters in deeper layers. Max pooling layers are applied periodically to reduce the spatial dimensions, which lowers computational demands while preserving essential features. The final three layers are fully connected, translating learned features into classification predictions. Although VGG16's simplicity aids its effectiveness, the architecture has a significant number of parameters, making it computationally demanding and memory-intensive. Additionally, it lacks residual connections, which can lead to vanishing gradient issues in very deep networks. Despite these limitations, VGG16 remains popular for transfer learning, as pretrained weights on large datasets like ImageNet can be adapted to various computer vision applications. Its balance of simplicity and performance makes VGG16 a valuable architecture in the deep learning community, especially when computational resources are available.

Training the Model:

- To train the models, the training module first loads the pre-trained weights for MobileNetV2 and VGG16 from the dataset. It then adds a global average pooling layer and a fully connected layer with the appropriate number of classes for the specific task. The base layers of the pre-trained models are frozen by setting their trainable flag to False, allowing only the added layers to be trained.
- The training process involves optimizing the model parameters using a suitable optimization algorithm, such as Adam or SGD, and a loss function appropriate for the task (e.g., categorical cross-entropy for classification). The training data is fed to the model in batches, and the gradients are computed and used to update the model weights. The training process continues for a specified number of epochs or until a certain performance metric is achieved on a validation set.
- After training, the module saves the trained models for future use in the prediction and evaluation modules. The saved models can be loaded and used for

inference on new data or fine-tuned on additional datasets if needed.

Analyze and Prediction:

- Once training is complete, analyze the training process (e.g., loss curves) and make predictions on your validation set to assess the model's performance.

Accuracy on test set:

- Once the model is trained, it needs to be evaluated for its performance. This module involves splitting the dataset into training and testing subsets and assessing the model's accuracy, precision, recall, and F1-score.
- The MobileNetV2 architecture achieves a Training accuracy of 96.00% and validation accuracy of 96.00%. The VGG16 architecture attains a Training accuracy of 93.00% and validation accuracy of 92.00%.

Saving the Trained Model:

- Once you're confident enough to take your trained and tested model into the production-ready environment, the first step is to save it into an .h5 or .pkl file using a library like pickle.
- Make sure you have pickle installed in your environment.
- Next, let's import the module and dump the model into .pkl file.

Prediction Module:

- Develop a prediction module to make predictions using the trained MobileNetV2 and VGG16 model. This module should take input images; preprocess them as necessary, and output predictions.

Model Evaluation Module

- This module evaluates the performance of the trained models using the testing dataset. It calculates accuracy metrics and other performance indicators to assess model effectiveness.
- Evaluate model accuracy, precision, recall, and F1-score.
- Generate confusion matrices for both models.
- Compare the performance of the MobileNetV2 and VGG16.
- Accuracy, precision, recall, and F1-score are used to evaluate model performance.

CONCLUSION

The project "Automated Knee Osteoarthritis Prediction and Classification from X-ray Images Using Deep Learning" successfully demonstrates the potential of deep learning in automating the diagnosis and grading of knee osteoarthritis. By leveraging state-of-the-art models like VGG16 and MobileNetV2, the system achieves high accuracy, with the MobileNetV2 model excelling at a test accuracy of 96%.

The integration of advanced preprocessing techniques, including novel cartilage region extraction, ensures that the models focus on the most relevant features of the X-ray images, improving diagnostic precision. The carefully curated dataset, annotated by medical experts, adds reliability and robustness to the system's predictions. Moreover, the comprehensive performance evaluation and user-friendly web interface make the system suitable for practical applications in clinical settings.

This project bridges the gap between technological advancements and healthcare needs, providing a scalable, accurate, and efficient solution for detecting and grading knee osteoarthritis. The results highlight the significant impact of deep learning in medical imaging and underline its utility in enhancing diagnostic workflows and supporting healthcare professionals.

FUTURE WORK:

- Integration of Multimodal Data: Future enhancements could include integrating additional data modalities, such as patient demographics, medical history, and genetic information, alongside X-ray images to improve the model's diagnostic accuracy and provide a more holistic assessment of knee osteoarthritis.
- Real-Time Prediction Capability: Implementing real-time processing capabilities for live X-ray scans could enhance the system's practicality in clinical settings, enabling instantaneous predictions and reducing diagnostic delays.
- Deployment on Mobile and Edge Devices: Optimization of the system for deployment on mobile devices and edge computing platforms would make it accessible to remote or resource-constrained areas, increasing its impact and usability.
- Incorporation of Advanced Explainability Tools: Adding model explainability features, such as Grad-CAM or SHAP values, could help clinicians understand the basis of the predictions, increasing trust in the system and aiding in decision-making.
- Support for Additional Imaging Modalities: Expanding the system to support other imaging modalities, such as MRI or CT scans, could broaden its application scope and improve diagnostic capabilities for complex cases of osteoarthritis.

- Dataset Expansion and Diversity: Enhancing the dataset with images from diverse sources, including various age groups, ethnicities, and medical conditions, would improve the system's generalizability and robustness across a wider patient population.
- Integration into Electronic Medical Record (EMR) Systems: Developing an interface for seamless integration with existing EMR systems could enable healthcare providers to access and utilize predictions directly within their workflow, enhancing efficiency and patient care.
- Longitudinal Monitoring: Incorporating functionality for tracking the progression of knee osteoarthritis over time using serial X-ray images could help clinicians evaluate the effectiveness of treatment plans and make informed decisions.
- Multi-Class Classification Refinement: Further refinement of the multi-class classification, especially for borderline cases, could improve the accuracy and reliability of grading across the Kellgren and Lawrence scale.
- Clinical Trials and Validation: Conducting extensive clinical trials to validate the system's effectiveness in real-world scenarios and obtaining regulatory approvals would be a key step toward its adoption in healthcare practice.
- By addressing these future directions, the system can evolve into a more comprehensive and impactful tool for knee osteoarthritis diagnosis and management.

REFERENCES

[1] A. Cui, H. Li, D. Wang, J. Zhong, Y. Chen, and H. Lu, "Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies," EClinicalMedicine, vols. 29–30, Dec. 2020, Art. no. 100587, doi: 10.1016/j.eclinm.2020.100587.

[2] P. P. F. M. Kuijer, H. F. van der Molen, and S. Visser, "A health-impact assessment of an ergonomic measure to reduce the risk of work-related lower back pain, lumbosacral radicular syndrome and knee osteoarthritis among floor layers in The Netherlands," Int. J. Environ. Res. Public Health, vol. 20, no. 5, p. 4672, Mar. 2023, doi: 10.3390/ijerph20054672.

[3] V. K. V, V. Kalpana, and G. H. Kumar, "Evaluating the efficacy of deep learning models for knee osteoarthritis prediction based on kellgrenlawrence grading system," e-Prime - Adv. Electr. Eng., Electron. Energy, vol. 5, Sep. 2023, Art. no. 100266, doi: 10.1016/j.prime.2023.100266.

[4] S. Castagno, M. Birch, M. van der Schaar, and A. McCaskie. (2024). Prediction of the Rapid Progression of Knee Osteoarthritis Using Automated Machine Learning: A Novel Precision Health Approach for Chronic Degenerative Diseases. SSRN. Accessed: Aug. 2, 2024. [Online]. Available: https://ssrn.com/abstract=4561796

[5] A. A. S. Afroze, R. Tamilselvi, and M. G. P. Beham, "Machine learning based osteoarthritis detection methods in different imaging modalities: A review," Current Med. Imag. Formerly Current Med. Imag. Rev., vol. 19, no. 14, pp. 1628–1642, Dec. 2023, doi: 10.2174/1573405619666230130143020.

[6] Y. C. Park, K. J. Park, B. H. Goo, J. H. Kim, B. K. Seo, and Y. H. Baek, "Oriental medicine as collaborating treatments with conventional treatments for knee osteoarthritis: APRISMA-compliant systematic review and meta-analysis," Medicine, vol. 102, no. 29, 2023, Art. no. E34212, doi: 10.1097/MD.0000000000034212.

[7] S. V Chaugule, S. V Chaugule, and V. S. Malemath, "Towards identifying key features in the classification of knee osteoarthritis—An enhanced feature fusion based deep network model," Tech. Rep., 2023.

[8] Y. Soda, T. Kano, and M. Nakamura, "Kinematically aligned total knee arthroplasty for valgus osteoarthritis of more than 10◦ is it still a 'challenging surgery?'" Open J. Orthopedics, vol. 13, no. 9, pp. 355–369, 2023, doi: 10.4236/ojo.2023.139035.

[9] H. Khalid, M. Hussain, M. A. Al Ghamdi, T. Khalid, K. Khalid, M. A. Khan, K. Fatima, K. Masood, S. H. Almotiri, M. S. Farooq, and A. Ahmed, "A comparative systematic literature review on knee bone reports from MRI, X-rays and CT scans using deep learning and machine learning methodologies," Diagnostics, vol. 10, no. 8, p. 518, Jul. 2020, doi: 10.3390/diagnostics10080518.

[10] J. J. Bjerre-Bastos, M. A. Karsdal, M. Boesen, H. Bliddal, A. Bay-Jensen, J. R. Andersen, and A. R. Bihlet, "The acute and long-term impact of physical activity on biochemical markers and MRI measures in osteoarthritis—Perspectives for clinical osteoarthritis research," Transl. Sports Med., vol. 3, no. 5, pp. 384–394, Sep. 2020, doi: 10.1002/tsm2.155.

[11] Y. X. Teoh, A. Othmani, S. L. Goh, J. Usman, and K. W. Lai, "Deciphering knee osteoarthritis diagnostic features with

explainable artificial intelligence: A systematic review," 2023, arXiv:2308.09380.

[12] V. K. V, V. Kalpana, and G. H. Kumar, "Evaluating the efficacy of deep learning models for knee osteoarthritis prediction based on kellgrenlawrence grading system," e-Prime - Adv. Electr. Eng., Electron. Energy, vol. 5, Sep. 2023, Art. no. 100266.

[13] Y. Nasser, M. El Hassouni, D. Hans, and R. Jennane, "A discriminative shape-texture convolutional neural network for early diagnosis of knee osteoarthritis from X-ray images," Phys. Eng. Sci. Med., vol. 46, no. 2, pp. 827–837, Jun. 2023, doi: 10.1007/s13246-023-01256-1.

[14] S. Paul and P.-Y. Chen, "Vision transformers are robust learners," in Proc. AAAI Conf. Artif. Intell., 2022, vol. 36, no. 2, pp. 2071–2081, doi: 10.1609/aaai.v36i2.20103.

[15] K. Mohiuddin, M. A. Alam, M. M. Alam, P. Welke, M. Martin, J. Lehmann, and S. Vahdati, "Retention is all you need," in Proc. 32nd ACM Int. Conf. Inf. Knowl. Manage., Oct. 2023, pp. 4752–4758, doi:10.1145/3583780.3615497.