**A paper with text and words

AI-generated content may be incorrect.**

**1. Abstract**

Kidney diseases such as cysts, stones, and tumors are significant health concerns, often requiring timely and accurate diagnosis. In this midterm project, we focus on a Convolutional Neural Network (CNN) approach to classify kidney diseases using CT scan images. A curated dataset of kidney CT images was divided into four main classes: cyst, stone, tumor, and normal. After appropriate preprocessing, our CNN architecture was trained, achieving promising classification accuracy. These findings highlight the potential of CNNs in assisting medical professionals for automated kidney disease detection.

**2. Introduction**

Kidney-related conditions—including tumors, cysts, and stones—pose serious health risks if not diagnosed early. Traditionally, radiologists rely on CT scans for visual assessment, which is time-consuming and can vary based on individual expertise. Recently, **Deep Learning (DL)** methods, especially **Convolutional Neural Networks (CNNs)**, have demonstrated remarkable capabilities in medical image classification.

In this midterm report, we document our progress on building a **CNN-based classification system** for kidney diseases. Our scope is currently limited to the CNN methodology, with plans to extend or compare with other architectures in future work.

**3. Methodology**

**3.1 Dataset**

* **Source & Composition**: The dataset contains 12,446 CT images labeled into four categories:
  1. **Cyst**
  2. **Stone**
  3. **Tumor**
  4. **Normal**



**Stone Cyst Normal Tumor**

* **Data Split**: Training, Validation, and Testing splits were created to ensure balanced representation of each class.

**3.2 CNN Architecture**

Our CNN, referred to as **CNNModel**, is designed to extract features hierarchically from CT images:

1. **Convolutional & Pooling Blocks**
   * We used two convolutional blocks. Each block contains:
     + Two convolutional layers (kernel size 3×3)
     + Rectified Linear Unit (ReLU) activation
     + MaxPooling layer (stride 2) to reduce spatial dimensions
2. **Fully Connected (Dense) Layers**
   * After flattening the feature maps, two Dense layers are used for classification.
   * Dropout is applied in the Dense layers to mitigate overfitting.
   * The final Dense layer outputs 4 units (one for each class) using **Softmax** activation.
3. **Hyperparameters**
   * **Loss Function**: Cross-Entropy
   * **Optimizer**: Adam (learning rate = 1×10<sup>-3</sup>)
   * **Batch Size**: 10
   * **Epochs**: 32

A diagram of a diagram of a structure

AI-generated content may be incorrect.

**4. Experiments and Preliminary Results**

**4.1 Training Procedure**

* **Preprocessing**:
  + Images were resized to 224×224 pixels and normalized.
  + Data augmentation techniques (e.g., random rotations, flips) were optionally tested to reduce overfitting.
* **Model Training**:
  + The model was trained for up to 32 epochs, monitoring the **loss** and **accuracy** on both training and validation sets.

**4.2 Performance Metrics**

* **Accuracy**: Percentage of correctly classified images on the test set.
* **Confusion Matrix**: Counts of correct and incorrect predictions per class.

| **Model** | **Train Accuracy** | **Test Accuracy** |
| --- | --- | --- |
| **CNNModel** | ~90% | **~98%** |

*(Exact values may vary slightly depending on random seeds or hardware.)*

**4.3 Confusion Matrix (Illustrative)**

|  | **Predicted Cyst** | **Predicted Normal** | **Predicted Stone** | **Predicted Tumor** |
| --- | --- | --- | --- | --- |
| **Actual Cyst** | 99.2% | 0.4% | 0.2% | 0.2% |
| **Actual Normal** | 0.3% | 98.0% | 0.5% | 1.2% |
| **Actual Stone** | 0.6% | 0.6% | 98.3% | 0.5% |
| **Actual Tumor** | 0.4% | 1.0% | 0.2% | 98.4% |

The matrix demonstrates strong performance across all classes, with only minor misclassifications.

A diagram of a confusion matrix

AI-generated content may be incorrect.

**5. Discussion**

* **Key Observations:**
  + The CNN effectively discerns and assimilates distinct pathological features of kidney anomalies, resulting in commendable classification accuracy.
  + While the model demonstrates robust performance, the incorporation of dropout regularization was imperative to mitigate overfitting and enhance generalization.
  + Further optimization through meticulous hyperparameter tuning and extensive data augmentation holds the potential to fortify the model’s resilience and predictive robustness.
* **Challenges:**
  + The presence of class imbalance, where certain pathological categories are underrepresented, poses a constraint on model generalization.
  + Variability in medical imaging data—stemming from differences in equipment specifications, contrast settings, and acquisition protocols—introduces complexities in achieving consistent model performance across diverse datasets.

**6. Conclusion and Future Work**

Our CNN-based model for classifying kidney pathologies has shown promising preliminary results, with test accuracy reaching around 98%. This performance underscores the capability of CNNs to detect intricate visual patterns in CT scans.

**For future work, we plan to:**

As part of our ongoing research, we are actively working with **Vision Transformers (ViT) and pretrained models** to enhance model performance and interpretability. Our future directions include:

1. **Expanding Data Diversity** – Integrating additional medical imaging datasets from diverse sources to improve model generalization and robustness.
2. **Comparative Analysis** – Evaluating the performance of **Vision Transformers (ViT)** and other pretrained models against CNN-based approaches to determine their efficacy in kidney pathology classification.
3. **Ensemble Learning** – Developing a hybrid approach that synergizes CNN, ViT, and other pretrained architectures to optimize classification accuracy.
4. **Model Explainability** – Implementing advanced interpretability techniques (e.g., Grad-CAM, attention visualization in ViT) to provide deeper insights into model decision-making, aiding clinical applications.

**References**

1. Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., & Summers, R. M. (2016). *Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics, and transfer learning*. IEEE Transactions on Medical Imaging, 35(5), 1285-1298.
2. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
3. Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). *Convolutional neural networks for medical image analysis: Full training or fine-tuning?* IEEE Transactions on Medical Imaging, 35(5), 1299-1312.
4. Zhang, J., Shangguan, Z., Gong, W., & Cheng, Y. (2023). *A novel denoising method for low-dose CT images based on transformer and CNN*. Computers in Biology and Medicine, 163, 107162.