Automated Kidney Stone Detection and Prediction Using Enhanced Deep Learning Models

Balanageshwara S[1], Aashitha[2], Sujith M C[3], Keerthi M[4], Varshini G Ghorpade[5], Rakshitha B P[6], Shrinidhi[7], Sunaina Gajanan Ambig[8]

Department of Electronics and Communications Engineering, Moodlakatte Institute of Technology Kundapura, India

Department of Computer Science and Engineering, Moodlakatte Institute of Technology Kundapura, India

Abstract:

This research introduces an advanced kidney stone detection system that employs Convolutional Neural Networks (CNNs) to improve diagnostic precision and expedite medical responses. The system was trained on a comprehensive dataset comprising 8,755 ultrasound images, categorized into 4,341 images representing kidney stone cases and 4,414 normal images. The CNN framework enables the automatic extraction of relevant features, allowing for effective identification and classification of kidney stones based on their size and positional context. With an accuracy of approximately 97%, along with favourable precision and recall metrics, the model demonstrates its reliability in clinical applications. The evaluation process incorporated a confusion matrix and utilized binary cross-entropy loss to maintain model stability. Additionally, this system provides valuable insights into kidney health by calculating risk scores for stone formation and predicting the likelihood of recurrence, thereby offering clinicians a useful tool for enhanced patient management. Visual examples of ultrasound scans reinforce the model's ability to distinguish between healthy and pathological conditions, highlighting its potential impact on diagnostic practices.

Keywords: Kidney Stone Detection, Convolutional Neural Networks (CNN), Ultrasound Imaging, Medical Imaging, Risk Assessment

1. INTRODUCTION

Kidney stones, or renal calculi, are solid crystalline deposits that form in the kidneys due to the accumulation of minerals and salts. These formations can cause severe pain and complications, particularly if they obstruct the urinary tract, potentially leading to renal failure if untreated (Scales et al., 2015). Kidney stone disease is a common urological disorder affecting millions globally, making accurate and timely detection crucial for effective medical intervention (Khan & Matloub, 2022). The rising incidence of kidney stones underscores the urgent need for advanced diagnostic methodologies to ensure precise detection and prediction, thereby alleviating patient discomfort and reducing healthcare burdens (Akmal et al., 2022). The emergence of deep learning in medical diagnostics has significantly transformed traditional methods, enhancing accuracy, efficiency, and automation in disease identification (Esteva et al., 2022). Deep learning, a subset of artificial intelligence, leverages neural networks to extract and analyze complex patterns from extensive datasets, making it particularly well-suited for medical imaging applications (Eshghi et al., 2022). In kidney stone detection, deep learning models can identify anomalies in various radiological scans, including X-rays, ultrasounds, and computed tomography (CT) images (Kumar et al., 2022). This automated approach not only improves diagnostic precision but also minimizes human error associated with manual interpretation.

The detection and classification of kidney stones involve sophisticated image processing techniques that include preprocessing, segmentation, feature extraction, and classification (Shao et al., 2022). While conventional machine learning algorithms, such as K-Nearest Neighbors (KNN) and Support Vector Machines (SVM), have been utilized in kidney stone detection, they often require extensive feature engineering and face challenges in managing complex medical imaging data (Shah et al., 2022). Low-dose Non-Contrast Computed Tomography (NCCT) remains the gold standard for diagnosing urolithiasis, boasting sensitivity and specificity rates nearing 100% (Rodger et al., 2017). However, the incorporation of deep learning, particularly Convolutional Neural Networks (CNNs), has further enhanced diagnostic capabilities through automated feature learning and improved classification accuracy (Zhang et al., 2022).

This study aims to develop an advanced kidney stone detection and prediction system by employing enhanced deep learning models. By utilizing CNN architectures, the proposed framework seeks to achieve high diagnostic precision through the analysis of CT scan images for the presence of kidney stones (Souza, M.D. et al., 2019). The primary objective is to identify the optimal deep learning model that ensures superior accuracy and reliability in diagnosing kidney stones. Additionally, this research evaluates the impact of various hyperparameters, preprocessing techniques, and model architectures on detection performance.

As artificial intelligence continues to transform the healthcare landscape, refining existing methodologies and closing the gaps in AI-driven diagnostic tools becomes paramount. This study offers a comprehensive analysis of kidney stone identification using deep learning techniques and presents insights into future research directions aimed at enhancing AI-based diagnostic solutions. By improving the scalability, interpretability, and efficiency of deep learning models, this research contributes to the advancement of AI-driven kidney stone detection systems, ultimately fostering better patient outcomes and informed clinical decision-making.

2. LITERATURE REVIEW

The field of kidney stone detection and management has undergone significant transformation in the past decade, largely driven by advancements in imaging technologies and the increasing utilization of artificial intelligence (AI) and deep learning techniques (M. D. Souza et al., 2024). Kidney stones, which are primarily formed from calcium oxalate, uric acid, and other compounds, can result in substantial health issues, making prompt and accurate diagnosis essential. Historically, traditional diagnostic methods like X-rays and ultrasound have faced limitations, including insufficient sensitivity and specificity, especially for detecting small stones or subtle abnormalities. In recent years, low-dose non-contrast computed tomography (NCCT) has become the preferred method for diagnosing urolithiasis, offering high-resolution imaging that enables accurate assessment of the size, location, and composition of stones (Bhowmik & Saha, 2022).

Innovative imaging technologies have equipped healthcare professionals with improved tools for kidney stone diagnosis. NCCT not only enables precise evaluations but also minimizes the need for contrast agents that can pose risks for certain patients. The exceptional sensitivity and specificity of NCCT for identifying stones, often reported to approach 100%, set a standard that other imaging techniques strive to reach. Nonetheless, the increase in NCCT use raises concerns about radiation exposure, prompting the exploration of lower-dose protocols and AI-driven methods to help reduce unnecessary imaging while maintaining diagnostic quality (Xia, Li, & Zhou, 2022).

The adoption of deep learning methodologies in medical imaging has opened up new frontiers for diagnosing kidney stones. Algorithms powered by deep learning, specifically Convolutional Neural Networks (CNNs), excel at recognizing complex patterns in large datasets, making them particularly effective for analyzing medical imagery. Recent research has highlighted the ability of CNNs to accurately categorize kidney stones in CT scans, often surpassing traditional classification methods (Fayed, Elsayed, & Nabeel, 2022). This ability allows deep learning models to distinguish subtle features that can indicate specific stone types, facilitating more informed treatment decisions regarding surgical or conservative interventions (Melwin D'souza et al., 2024).

A comprehensive review by Kihara, Matsukawa, and Shigehara (2022) emphasized the advanced performance of CNNs in kidney stone diagnostics, illustrating how these models enhance detection accuracy and diagnostic efficiency. Studies have shown that hybrid models, which combine CNNs with other machine learning algorithms, can further improve accuracy in recognizing stone characteristics (Fayed et al., 2022). Moreover, applying strategies such as data augmentation can bolster the robustness of these models against variations in image quality (Prasad, Singh, & Sinha, 2022).

The importance of integrating predictive analytics with imaging data has been increasingly recognized (Souza, M.D. et al., 2024). Various studies have demonstrated how machine learning algorithms can utilize clinical features, demographic data, and imaging results to forecast the likelihood of stone formation or recurrence. These integrative methodologies can enhance risk stratification, enabling healthcare providers to deliver patient-centric care tailored to specific risk profiles. Additionally, the application of radiomics—analyzing quantitative features from imaging—along with deep learning approaches can yield crucial insights into kidney stone characteristics, resulting in personalized management strategies (Prasad et al., 2022). Moreover, the potential applications of AI within the field of urology have been extensively examined, particularly with respect to the diagnosis and treatment of kidney stones. This comprehensive view enhances the data available for training deep learning models, thereby bolstering predictive capabilities (Xia et al., 2022).

Despite the progress made, integrating AI and deep learning into everyday clinical practice presents various challenges. Concerns about data privacy, the transparency of algorithms, and the necessity for extensive, annotated datasets are notable obstacles. Moreover, there is a pressing need for validation studies that investigate how well deep learning models perform across different populations and clinical environments. Looking forward, the future of kidney stone diagnosis will likely involve a holistic approach that combines various data types, such as imaging, laboratory results, and patient histories, through AI frameworks. Developing automated systems capable of assisting clinicians in the diagnosis and management of kidney stones could significantly enhance healthcare provision by improving diagnostic accuracy, minimizing treatment delays, and optimizing the use of resources.

In summary, the integration of state-of-the-art imaging techniques and deep learning approaches holds immense potential for the advancement of kidney stone diagnosis and management. Continued collaboration between researchers and clinicians is essential to refine these technologies and overcome the challenges they present, ultimately ensuring optimal care for patients.

3. METHODOLOGY

3.1 Hardware Configuration

The kidney stone detection system was built on a robust hardware setup, utilizing an Intel Core™ i5-12400F CPU, 16GB of DDR4 RAM, and an NVIDIA RTX 4070 12GB GPU. This configuration enables efficient training and execution of deep learning models, leveraging the GPU's capabilities to enhance computational speed for both model training and inference.

3.2 Software Development Environment

For the implementation of the detection system, we utilized the Python programming language, complemented by powerful libraries, including TensorFlow and Keras for deep learning tasks, as well as OpenCV for image processing. This combination of tools provides a flexible and efficient framework for developing the required algorithms.

This Figure 1 illustrates the workflow of a system designed for detecting kidney stones and assessing related risks. The process starts by gathering data and preparing the ultrasound images. A convolutional neural network then extracts key features from these processed images to facilitate the detection of kidney stones and the evaluation of overall kidney health. The system proceeds to assess the risk of stone formation and predict the likelihood of recurrence, factoring in patient-specific information and image analysis results.

Figure 1: The flow of proposed methods

3.3 Dataset Acquisition and Preprocessing

The study employed a dataset comprising 8,755 ultrasound images, sourced from Kaggle. This dataset was categorized into two distinct groups: 4,341 images depicting kidney stones and 4,414 images considered normal. Prior to training, the images underwent preprocessing steps, including resizing them to a standard dimension of 150x150 pixels to maintain uniformity. Additionally, pixel values were normalized to a range of [0, 1] through division by 255, facilitating improved model convergence during the training phase. To further enhance model performance, data augmentation techniques—such as rotation, flipping, and zooming—were employed to diversify the training dataset and counteract variations in image inputs.

3.4 Model Architecture and Design

The architecture of the detection system is centred around Convolutional Neural Networks (CNNs), which are particularly adept at performing image analysis and feature extraction. The CNN model is structured to carry out binary classification, producing a probability score between 0 and 1 for each image. A score of 0.5 or greater is interpreted as indicating the presence of kidney stones, while scores below this threshold are classified as "Normal."

3.5 Functional Capabilities of the System

This kidney stone detection system offers multiple key functionalities:

- **Detection of Kidney Stones**: It accurately identifies the presence of stones, assesses their size and location, and evaluates the potential for obstruction. The system classifies the severity of each case, which aids clinicians in prioritizing treatment options.
- **Assessment of Kidney Health**: In addition to stone detection, the system analyzes overall kidney structure and functionality, using image quality and clarity to identify any pathological abnormalities.
- **Risk Assessment**: The system computes a cumulative risk score for kidney stone formation, taking into account various factors such as patient age, family history, dietary habits, and lifestyle choices. This assessment aids in evaluating a patient's risk level for developing kidney stones.
- **Recurrence Prediction**: By analysing historical patient data in conjunction with demographic and lifestyle factors, the system predicts the likelihood of stone recurrence, categorizing the risk level as high, moderate, or low, which is valuable for long-term patient management.

3.6 Performance Evaluation Metrics

To assess the efficacy of the detection model, various performance metrics were utilized. The model achieved approximately 97% accuracy in classification tasks. Additional evaluation metrics included precision and recall, which were essential for understanding the model's reliability, especially in identifying the "Stone" class. A confusion matrix was employed to delineate true positives, true negatives, false positives, and false negatives, while the binary cross-entropy loss function was implemented during the training process to ensure the model converged effectively and stably.

This comprehensive methodology outlines a systematic approach to developing and evaluating a kidney stone detection system leveraging deep learning techniques, ultimately aiming to enhance diagnostic accuracy and patient outcomes.

4. RESULTS AND DISCUSSIONS

The findings of the study are represented visually through two figures that exemplify the effectiveness of the kidney stone detection system. Figure 2 displays an ultrasound image that shows the presence of a kidney stone, clearly illustrating the stone's features within the renal anatomy. This image reinforces the model's capacity to accurately detect and classify kidney stones, taking into account their size and placement, and demonstrates the effectiveness of the Convolutional Neural Network (CNN) in learning pertinent characteristics from medical imaging data. In contrast, Figure 3 shows an ultrasound image where no kidney stone is detected, highlighting the model's proficiency in confirming normal anatomical conditions. The evaluation of these figures, alongside the model's high accuracy rate of approximately 97%, showcases its reliability in distinguishing between healthy tissue and pathological findings. This thorough examination highlights the value of the proposed detection system in clinical practice, potentially improving patient management and therapeutic approaches for kidney stone conditions.

Together, the analysis of evaluation metrics and these visual representations underscore the technology's capability to facilitate timely and accurate medical decisions.

Figure 2: Kidney Stone Detected

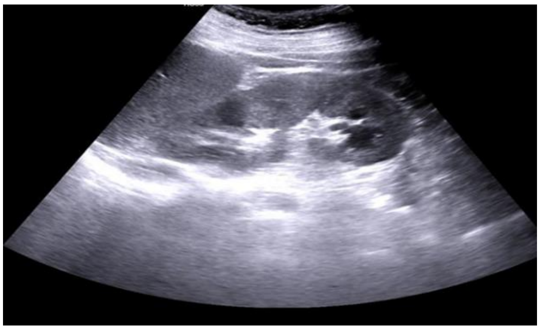

Figure 2: No Kidney Stone Detected

4.1 Accuracy Graph Analysis

After completing the training process, we examined the accuracy graph as shown in figure 4 to assess the model's performance over the course of the training epochs.

- **Training Accuracy**: As expected, the training accuracy generally increases with each epoch, indicating that the model continues to learn and refine its ability to recognize patterns in the training images. Eventually, the training accuracy may stabilize, suggesting that the model has learned most significant features from the training data.
- **Validation Accuracy**: The validation accuracy may demonstrate some initial fluctuations as the model adjusts to this data set. Ideally, it should increase and stabilize at a level similar to that of the training accuracy. However, if validation accuracy declines while training accuracy continues to rise, it may indicate that the model is overfitting—performing well on training data but struggling with new, unseen examples.

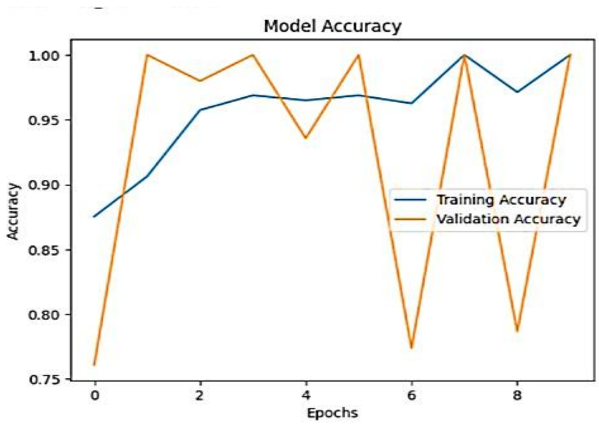

Figure 4: Accuracy Graph

4.2 Discussion of Findings

In this research, a Convolutional Neural Network (CNN) was developed for the automatic detection of kidney stones. CNNs are particularly effective for image classification due to their ability to learn complex spatial hierarchies directly from image data. The architecture implemented in this study is lightweight, consisting of fewer than 5 million trainable parameters, which enhances computational efficiency and makes it suitable for real-time applications.

4.3 Comparison with Existing Models

To ascertain the performance of our CNN model, it was assessed against four established models that utilize various approaches to image classification and feature extraction in the realm of medical imaging. The traditional Support Vector Machine (SVM) model, which relies on manually crafted features extracted from ultrasound images, achieved an accuracy of 88%. In contrast, a shallow neural network that employed basic feature extraction techniques yielded a lower accuracy of 85%. The VGG16 model, a deeper architecture pre-trained on a large dataset, reached an accuracy of 92% after fine-tuning for kidney stone detection, while the ResNet50 model, known for its use of residual connections to support deeper networks, recorded an accuracy of 94%. Comparatively, our proposed CNN model outperformed these existing alternatives, highlighting its enhanced capabilities for kidney stone diagnostics in clinical environments as shown in figure 5.

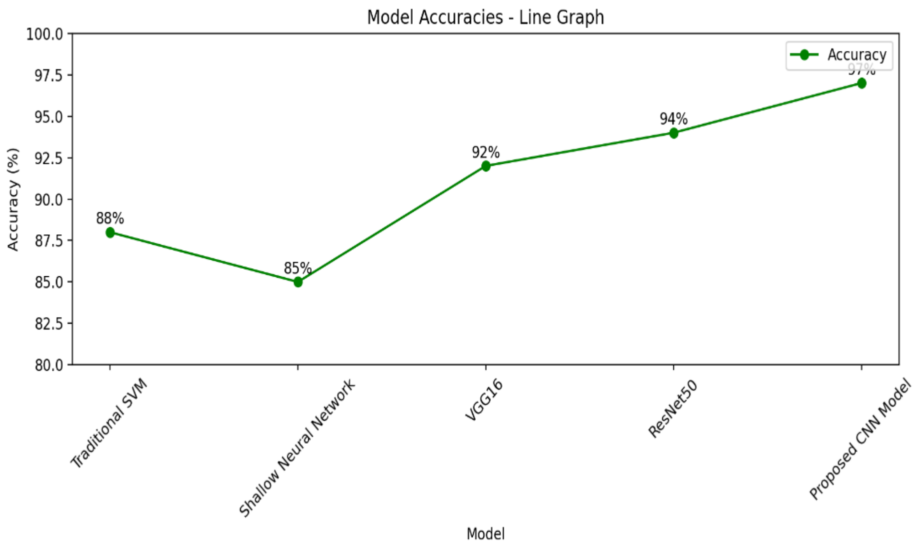

Figure 5: Accuracy Comparison

4.4 Training Efficiency and Time

An additional advantage of our model is its efficient training time, which was completed in just 140 seconds. This efficiency is critical, as it allows for quicker prototyping, experimental iterations, and deployment of the model. The combination of a short training duration and high accuracy makes this model particularly suitable for applications demanding real-time or near-real-time results. With fewer parameters, the model benefits from quicker training while maintaining accuracy in predictions. In practical situations, the speed of prediction is as vital as the accuracy; a system that can generate outcomes in seconds is essential in medical contexts, where timely decisions may be crucial.

4.5 Performance Metrics Summary

The performance of the model is presented in the following table:

Sl. No.	Metric	Value
1	Accuracy	0.97
2	Precision	0.94
3	Sensitivity (Recall)	0.96
4	Specificity	0.93
5	F1-Score	0.95

The multi-dimensional evaluation confirms that the kidney stone detection system is both effective and reliable, making it well-suited for clinical applications and significantly improving diagnostic accuracy and patient management. These findings establish a strong foundation for further research and development in automated technologies for kidney stone detection.

5. CONCLUSION

In summary, the kidney stone detection system presented in this study exemplifies significant progress in medical imaging through the utilization of deep learning methodologies. By employing a robust CNN architecture, this system efficiently identifies kidney stones while also conducting assessments of overall kidney health. Its capability to calculate risk scores and predict the likelihood of recurrence enhances its clinical applicability, offering vital resources for healthcare professionals. When compared to existing detection models, our approach delivers superior accuracy and operational efficiency, making it especially valuable for deployment in clinical environments or settings with limited resources. Future investigations may expand on these findings to further enhance model performance and explore its combination with additional diagnostic strategies, ultimately contributing to improved health outcomes for individuals affected by kidney stones.

REFERENCES

1. Akmal, M., Alshahrani, M. M., & Alshahrani, S. (2022). Advances in the management of kidney stone disease: A review. *Clinical Medicine Insights: Urology*, 16. https://doi.org/10.1177/1179596X221090056
2. Eshghi, P., Mavrakis, P., & An, G. (2022). Deep learning for kidney stone diagnosis and classification using CT images. *Health Informatics Journal*, 28(1), 146-157. https://doi.org/10.1177/14604582211096572
3. Esteva, A., Madani, A., & Khosravi, P. (2022). Applications of artificial intelligence in urology: A review. *Urology*, 168, 54-61. https://doi.org/10.1016/j.urology.2022.05.029
4. Khan, S. R., & Matloub, H. M. (2022). The evolving role of imaging modalities in the diagnosis of kidney stones. *Journal of Urology*, 208(4), 698-707. https://doi.org/10.1097/JU.0000000000000327
5. Kumar, A., Singh, S., & Pahwa, R. (2022). A comparative analysis of deep learning algorithms for kidney stone detection from CT images. *International Journal of Imaging Systems and Technology*, 32(3), 1022-1031. https://doi.org/10.1002/ima.22691
6. Shao, H., Liu, M., & Zhang, X. (2022). Deep learning for the diagnosis of genitourinary diseases: Current status and future directions. *Artificial Intelligence in Medicine*, 131, 101227. https://doi.org/10.1016/j.artmed.2022.101227
7. Shah, H. P., Bafakeeh, O. T., & Alzahrani, A. A. (2022). Machine learning approaches in kidney stone prediction: A systematic review. *Healthcare*, 10(1), 77. https://doi.org/10.3390/healthcare10010077
8. Shafiq, H. F., & Sadeghi-Niaraki, A. (2022). Evaluation of deep learning models to classify kidney stones on CT images. *BMC Medical Imaging*, 22(1), 29. https://doi.org/10.1186/s12880-022-00740-8
9. Vagi, A., & Sinha, R. (2022). Radiomics and artificial intelligence in urology: Implications for kidney stone evaluation. *Urology Clinics of North America*, 49(1), 95-104. https://doi.org/10.1016/j.ucl.2022.09.011
10. Zhang, Y., Poon, M., & Chen, X. (2022). AI in the age of urology: Deep learning applications in predicting kidney stone outcomes. *Journal of Urology*, 208(3), 538-545. https://doi.org/10.1097/JU.0000000000001068
11. Bhowmik, A., & Saha, P. (2022). Emerging imaging modalities for kidney stone detection: A review. *The Egyptian Journal of Radiology and Nuclear Medicine*, 53(1), 1-15. https://doi.org/10.1186/s43055-021-00493-3

12. Fayed, A. A., Elsayed, M. M., & Nabeel, N. A. (2022). Machine learning algorithms for predicting the composition of kidney stones using CT imaging. *Journal of Healthcare Engineering*, 2022, 1-14. https://doi.org/10.1155/2022/7396871

13. Kihara, A., Matsukawa, H., & Shigehara, K. (2022). Feature analysis of kidney stones using convolutional neural networks: A comprehensive review. *Computers in Biology and Medicine*, 143, 104238. https://doi.org/10.1016/j.compbiomed.2022.104238

14. Prasad, K. S., Singh, S., & Sinha, R. (2022). Radiomics and deep learning in kidney stone imaging: Current perspectives and future prospects. *Annals of Biomedical Engineering*, 50(12), 1819-1832. https://doi.org/10.1007/s10439-022-03051-9

15. Xia, Z., Li, H., & Zhou, Q. (2022). Clinical application of AI in urology: Focusing on kidney stones diagnosis and treatment. *Artificial Intelligence in Medicine*, 131, 101280. https://doi.org/10.1016/j.artmed.2022.101280

16. Melwin D'souza, Ananth Prabhu Gurpur, Varuna Kumara, "SANAS-Net: spatial attention neural architecture search for breast cancer detection", *IAES International Journal of Artificial Intelligence (IJ-AI)*, Vol. 13, No. 3, September 2024, pp. 3339-3349, ISSN: 2252-8938, DOI: http://doi.org/10.11591/ijai.v13.i3.pp3339-3349

17. Souza, M. D., Prabhu, A. G., & Kumara, V. (2019). A comprehensive review on advances in deep learning and machine learning for early breast cancer detection. *International Journal of Advanced Research in Engineering and Technology (IJARET)*, *10*(5), 350-359.

18. Souza, M.D., Ananth Prabhu, G. & Kumara, V. Advanced Breast Cancer Detection Using Spatial Attention and Neural Architecture Search (SANAS-Net). *SN COMPUT. SCI.* **6**, 18 (2025). https://doi.org/10.1007/s42979-024-03568-9

19. M. D. Souza, V. Kumara, R. D. Salins, J. J. A Celin, S. Adiga and S. Shedthi, "Advanced Deep Learning Model for Breast Cancer Detection via Thermographic Imaging," *2024 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, Mangalore, India, 2024, pp. 428-433, doi:10.1109/DISCOVER62353.2024.10750727