HIGH SCHOOL EDITION
Journal of Student Research

Volume 13 Issue 3 (2024)

Safer Surgeries Using AI: Enhancing Surgical Accuracy with Machine Learning Algorithms for Instrument Counting

Zenia Haroon

Glenelg High School, USA

ABSTRACT

This study explores the integration of Convolutional Neural Networks (CNNs) into surgical procedures to address the critical issue of retained surgical instruments (RSIs), enhancing patient safety. RSIs are when a foreign object is unintentionally left within a patient after surgery or other invasive procedures. This paper discusses challenges associated with manual counting methods and the potential of CNNs in automating surgical tool recognition. It employs an experimental design to investigate the impact of model hyperparameters on machine learning performance, utilizing the "Labeled Surgical Tools and Images" dataset for training. Preprocessing techniques, data augmentation, and Dropout regularization enhanced model robustness. Results from multiple training trials demonstrate the efficacy of the CNN-based model in accurately identifying and classifying surgical instruments, even in limited data scenarios. This research identifies overfitting as a challenge and addresses it through model adjustments and regularization techniques; it also highlights the findings' implications for improving surgical instrument count accuracy and enhancing patient safety. This study concludes by emphasizing the transformative potential of CNNs in surgical practice and the importance of ongoing research further to advance machine learning technologies in real-world surgical settings.

Introduction

Surgical procedures are transforming by integrating cutting-edge technologies, particularly CNNs. The issue of surgical instrument count accuracy during surgeries is paramount to ensure patient safety. About a dozen sponges and other surgical instruments are left inside patients' bodies every day in the United States, with approximately 70% of the total items left being sponges and the remaining 30% being surgical instruments such as clamps and retractors (Gawande et al., 2003). Retained surgical instruments (RSIs) are common and are between 0.3 and 1.0 per 1,000 abdominal operations (Zejnullahu et al., 2017). Manual counting methods are prone to errors due to insufficient organization and communication between surgical staff, leading to severe complications (Gawande et al., 2003; Zejnullahu et al., 2017). In some cases, these retained objects can result in sepsis or death, as well as localized pain, discomfort, and bloating. This problem is entirely avoidable; machine learning algorithms can be utilized to improve this process.

In this paper, the author discusses the challenges associated with manual counting and machine learning and the ethical concerns surrounding data collection for such models, as surgical data cannot be collected or tested. It also discusses the expansive realm of CNN applications within surgical tool recognition, examining the intricacies of real-time identification and classification. This research aims to explore the development of a machine learning model optimized for limited data scenarios, particularly addressing ethical concerns related to data collection and utilization in the medical field, which can be solved using surgical simulators in the future. These simulators could serve as a viable resource for data generation and offer a controlled environment to fine-tune models, paving the way for advancements that ensure both efficacy and ethical integrity in surgical tool recognition. Convolutional Neural Networks offer a promising solution to enhance surgical instrument count accuracy during surgery, addressing a critical aspect of patient safety.

Literature Review

Surgical Tool Recognition and the Utilization of Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a class of deep learning algorithms designed to automatically and adaptively learn hierarchical features from data. Due to their ability to effectively process and analyze visual data, CNNs have gained widespread adoption in various fields, including medical imaging and surgical tool recognition (Jaafari et al., 2021). CNNs excel in detecting patterns and features within images, making them particularly well-suited for tasks such as surgical tool detection and classification. CNNs leverage layers of interconnected neurons that apply convolutional and pooling operations to input pictures, extracting relevant features at different spatial scales. These features are then passed through additional layers, such as fully connected layers, to generate predictions or classifications. The hierarchical nature of CNNs allows them to learn complex representations of surgical instruments, enabling accurate real-time identification and classification in real-life surgical settings (Lam et al., 2022).

Recognition of surgical tools and anatomical features in surgical video images is critical in modern surgical settings. Real-time detection and identification of instruments and anatomical structures facilitates surgical navigation, enhances surgical efficiency, and reduces the risk of errors. Bamba et al. (2021) studied automated recognition of objects and types of forceps in surgical images. Their findings showcased the effectiveness of CNNs in accurately identifying different kinds of forceps with high precision and recall rates. Similarly, Garcia-Peraza-Herrera et al. (2021) explored image compositing techniques for segmenting surgical tools without manual annotations, demonstrating the potential of CNN-based approaches in overcoming data annotation challenges and achieving accurate recognition results. Though neither of these papers specifically addresses the issue of surgical instrument counting, they use computer vision to detect surgical tools, showing that machine-learning models can solve the problem of RSIs.

Current Prevention of Retained Surgical Instruments (RSIs)

Currently, surgical teams, including Certified Surgical Technologists (CSTs) and circulators, regularly count sponges, sharps, and instruments throughout surgeries to prevent the retention of foreign objects. These counts are performed audibly and documented as accurately as possible to minimize the risk of errors. However, this is a manual and traditional method of counting surgical tools, and with machine learning, RSIs can be easily avoided. Additionally, facilities may implement policies and procedures to ensure proper handling and disposal of instruments and training programs to educate surgical team members on best practices for instrument management (Association of Surgical Technologists [AST], n.d.).

Weprin et al. (2021) emphasize the importance of implementing standardized protocols and systematic search methods during surgical procedures, which include thorough counting methods, technology such as Radio-frequency identification (RFID) tagging or magnetic retrieval devices, and systematic search protocols to locate lost items, particularly in minimally invasive surgeries with limited visual scope. Weprin et al. also advocate for computer-aided detection to assist surgical teams in preventing RSIs. Additionally, another recent paper was released in 2023 aiming to solve the problem of RSIs using machine learning, proposing hospitals integrate video technology with AI-powered cameras and analytic systems. These technologies would monitor surgical procedures in real time, track the use and disposal of surgical instruments, and minimize the risk of items being left inside patients. They propose that the open-platform video management software (VMS) enables object recognition, allowing surgical teams to cross-reference detected items with surgical checklists to ensure all instruments are accounted for before incisions are closed. AI-powered cameras integrated with video management alarm systems can detect missing or unusual objects, alerting the surgical team to rectify discrepancies immediately. Additionally, predictive analytics would analyze recorded video data to identify patterns that may lead to incidents of RSIs. This would enable hospitals to implement proactive measures and training programs to prevent such occurrences (Global Edition Artificial Intelligence, 2023).

Although the above-mentioned methods are believed to reduce RSIs, many surgeons are reluctant to allow their procedures to be recorded, refuting both of these sources' proposals. Surgeons may be wary of legal liability for any "accidents." Still, although it is an influential factor, they mainly abstain from recording surgeries due to significant privacy concerns. Recording surgeries violates the Health Insurance Portability and Accountability Act (HIPAA), a severe offense. While it is feasible and occasionally conducted for educational purposes, it necessitates explicit written consent from the patient. Many other paperwork must be signed before even looking at the recording, which is a tedious process to get approved. To address this concern, this paper proposes placing a camera over the sterile field tray instead of on top of the patient to address ethical concerns. The AI would then compare the list of required materials to the initial number of tools placed on the table and then again to the final amount of tools. If tools are missing, the AI will alert the medical staff to take action, providing the exact tool they need to look for.

Computer Aided Surgery (CAS)

Computer-Aided Surgery (CAS) represents a transformative approach to surgical practice, leveraging advanced technologies to enhance various aspects of surgical procedures. One of the critical areas of exploration within CAS is the application of computer vision technologies. These technologies offer a wide range of potential applications in surgical practice, including improved surgical education, enhanced navigation during procedures, and increased overall procedural efficiency (Lee et al., 2021). Exploring potential applications of computer vision technologies in computer-aided surgery encompasses various domains within surgical practice. One notable application is in surgical education, where computer vision systems can facilitate interactive learning experiences for medical students and practitioners. By providing real-time visualizations and simulations of surgical procedures, these systems offer invaluable training opportunities and enhance the understanding of complex surgical techniques (Chiew et al., 2019). Another significant application is in intraoperative navigation, where computer vision technologies enable surgeons to precisely locate anatomical structures and navigate surgical instruments with greater accuracy. By integrating real-time imaging data with preoperative scans and intraoperative observations, these systems enhance surgical precision and reduce the risk of procedural errors (Chinedu et al., 2022).

Challenges and Limitations

One of the primary challenges in surgical tool recognition is the availability and quality of annotated datasets for model training. While large-scale annotated datasets are essential for effectively training deep learning models, acquiring such datasets in the medical domain poses significant challenges, mainly due to HIPAA (Prellberg & Kramer, n.d.). Annotated surgical datasets are often scarce due to privacy concerns, limited access to surgical footage, and the expertise required for manual annotation. Moreover, the quality of annotations can vary, leading to inconsistencies and biases in the training data, which may affect model performance negatively. Addressing these challenges necessitates the development of standardized protocols for dataset collection, annotation, and curation, ensuring data quality and diversity to enhance model robustness.

Class imbalance and variability in real-world surgical scenarios pose additional challenges to surgical tool recognition systems. In surgical environments, specific tools may be used more frequently than others, leading to imbalanced class distributions in the training data (Garcia-Peraza-Herrera et al., 2021). This imbalance can bias the model towards the dominant classes, reducing performance for minority classes. Furthermore, the variability in surgical procedures, lighting conditions, camera perspectives, and patient anatomies introduces additional complexity to the recognition task. Models trained on limited or homogeneous datasets may struggle to generalize to diverse surgical settings, highlighting the importance of dataset diversity and augmentation techniques (Chen et al., 2022). Addressing these complexities requires interdisciplinary collaborations between computer scientists, medical professionals, and ethicists to develop robust, interpretable, and clinically relevant recognition systems. This is why the author of this paper proposes using surgical simulation, such as the videogame *Surgical Simulator 2*, to train and base the image classification model, which will expand the dataset. As the project progresses, the model can be applied to physical scenarios, possibly using a Raspberry Pi and a camera to detect surgical tools in real-time.

Methods

For this research, an experimental design was employed, utilizing a quantitative approach to investigate the impact of model hyperparameters on the performance of various machine learning models. This design involves a randomized controlled trial where the independent variable, model hyperparameters, were systematically manipulated to observe their effects on the dependent variables: model accuracy and precision. These metrics serve as crucial benchmarks for evaluating the effectiveness of machine learning models in real-world applications. This research utilizes the Kaggle dataset "Labeled Surgical Tools and Images" by Diana Lavado, chosen for its suitability in studying the impact of model hyperparameters on machine learning performance, as it has clean and simple data.

Images underwent preprocessing steps such as resizing, normalization, and augmentation to ensure uniformity in size, quality, and variability in the dataset—preprocessing aimed to enhance model performance and generalization by minimizing overfitting. Dataset generation and augmentation are vital for training neural networks in surgical tool recognition. Various methods, like collecting surgical videos and applying augmentation techniques, enhance dataset diversity and quality (Bamba et al., 2021; Lehr et al., 2023). However, data scarcity and the need for expert annotation hinder dataset creation. Generative Adversarial Networks (GANs) offer a solution by synthesizing realistic medical images, addressing data scarcity in laparoscopic image analysis (Marzullo et al., 2021; Praveen SR Konduri & Siva, 2024). Diverse synthetic datasets are crucial for improving model robustness and clinical applicability, capturing the variability of surgical procedures and patient anatomies (Lee et al., 2021).

The choice of a CNN architecture was motivated by its effectiveness in image classification tasks. A sequential model architecture was employed, consisting of convolutional layers for feature extraction and dense layers for classification. The CNN model architecture was configured with appropriate input dimensions, convolutional layer parameters (filter size, number of filters), activation functions, and dropout layers to mitigate overfitting. The dataset was divided into training and validation sets to evaluate the model performance during training. The training set was used to optimize model parameters, while the validation set facilitated monitoring of model generalization. The author also manually cleaned the data by selecting only images where each surgical tool was alone and not with other tools to create a simplistic baseline version as a test bed to conduct the study. Hyperparameters such as learning rate, batch size, and number of epochs were tuned to optimize model convergence and performance. The model was trained using an iterative process, where training examples were presented to the model in batches. During training, the model learned to minimize a predefined loss function called categorical cross-entropy by adjusting its parameters using optimization algorithms like Adam. The performance of the trained model was evaluated using standard evaluation metrics such as accuracy, precision, recall, and F1-score on both the training and validation sets. These metrics provided insights into the model's classification capabilities and potential areas for improvement. Line graphs were implemented to illustrate how specific model hyperparameters influence accuracy and precision based on the data. They were also crucial in assessing model behavior, identifying misclassifications, and diagnosing potential issues such as overfitting or underfitting. These visualizations enhance the interpretation of the results and facilitate a deeper understanding of the underlying trends and patterns.

The image classifier was developed and trained using TensorFlow and relevant Python libraries for machine learning and image processing. The experiments were performed on computing hardware with sufficient processing power, memory, and GPU acceleration to facilitate efficient model training and evaluation.

Data

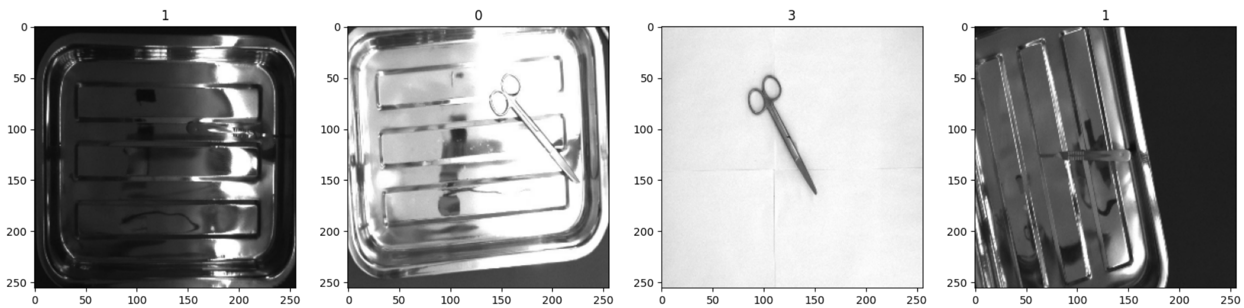

Figure 1. Sample scaled images from the "Labeled Surgical Tools and Images" dataset

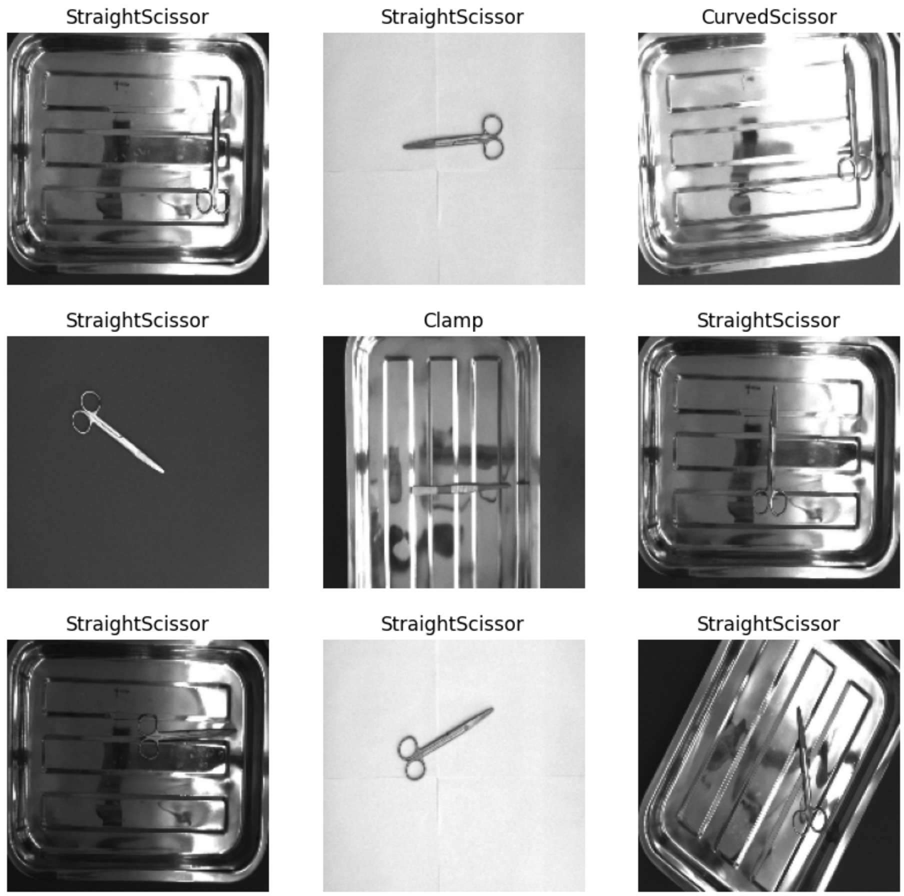

Figure 2. Sample scaled images from the "Labeled Surgical Tools and Images" dataset

The dataset used is the "Labeled Surgical Tools and Images" by Diana Lavado. The collection consists of 3009 images and labels that classify the objects as Scalpels, Straight Dissection Clamps, Straight Mayo Scissors, or Curved Mayo Scissors, and where each tool resides (on top, not occluded, or at the bottom, occluded). To allow the image classifier to recognize individual tools better, the author utilized 450 alone images of the straight mayo scissor, 460 alone images of the straight dissection clamp, 550 alone images of the scalpel, and 550 alone images of the curved mayo scissor. Sample scaled data are shown in Figures 1 and 2.

Trial 1

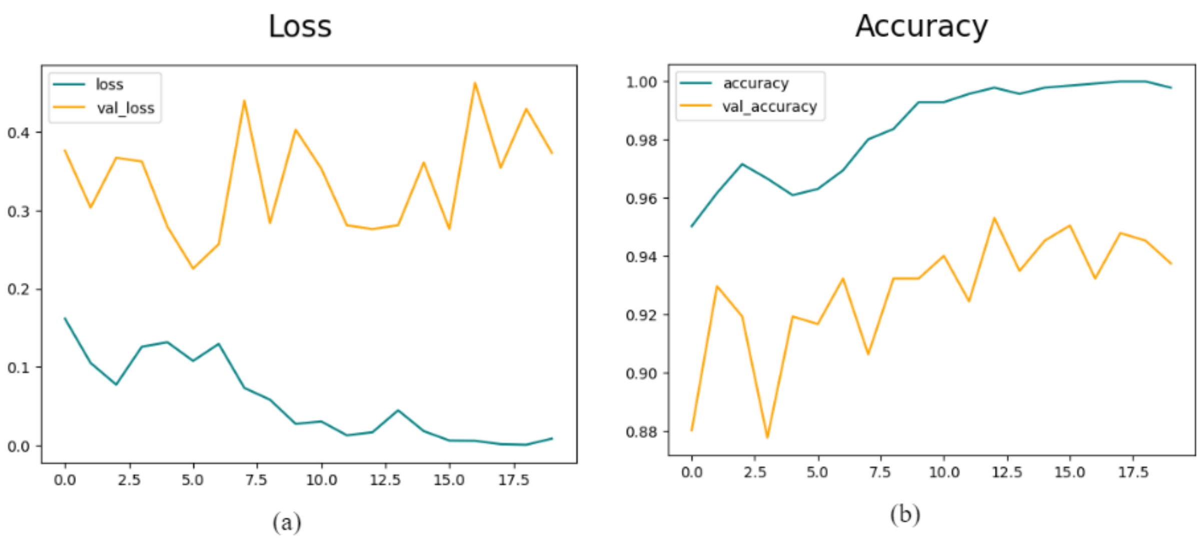

Figure 3. (a) Loss graph of the first training data of the surgical tools image classifier. (b) Accuracy graph of the first training data of the surgical tools image classifier

Loss measures the model's performance, with lower values indicating better performance. Accuracy shows the proportion of the training data the model classified correctly. Validation loss means the loss value on a separate validation dataset used to evaluate the model's performance on unseen data. Validation accuracy represents the model's accuracy on the validation data, which provides insight into its generalization capabilities. After training for 20 epochs, the loss was 0.0089, the accuracy was 0.9979, the validation loss was 0.3730, and the validation accuracy was 0.9375, as shown in Figure 3.

Trial 2

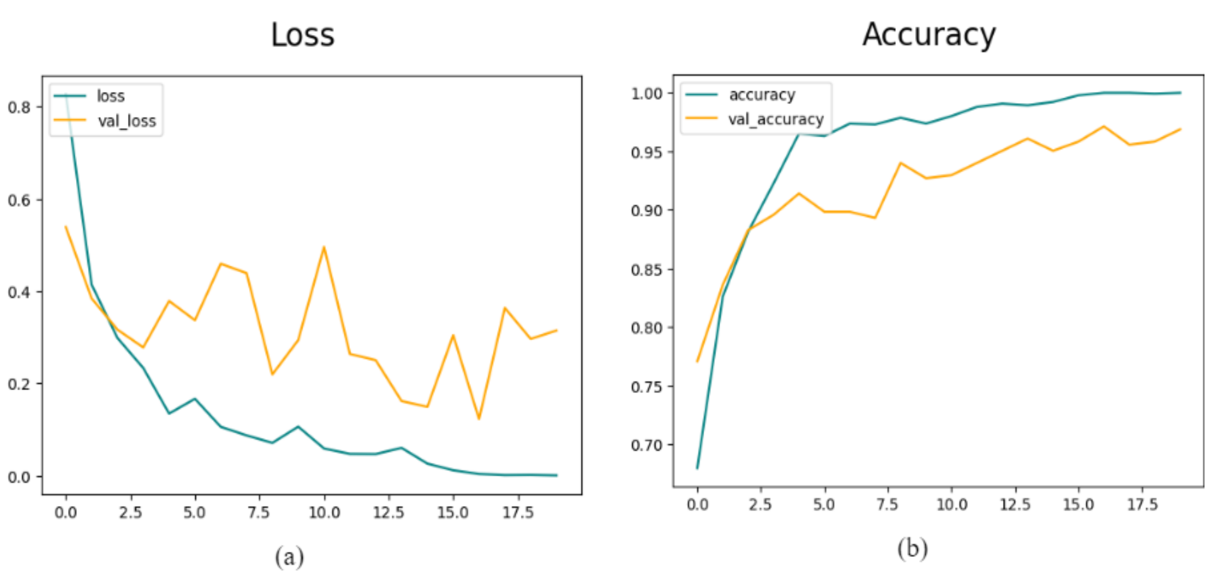

Figure 4. (a) Loss graph of the second training data of the surgical tools image classifier. (b) Accuracy graph of the second training data of the surgical tools image classifier

To ensure that human error is not a factor and to confirm that the model is genuinely not performing as well as had hoped, the author then trained the machine learning model a second time, again for 20 epochs. The loss was 6.3671e-04, the accuracy was 1.0000, the validation loss was 0.3147, and the validation accuracy was 0.9688. However, although the accuracy was high, the model failed to generalize well to unseen data. This situation, overfitting, occurs when the model learns to memorize the training data rather than capture the underlying patterns. As a result, the model performs poorly on new, unseen data despite its high accuracy on the training set. In Figure 4, although the training loss and accuracy decreased and increased, respectively, the validation fluctuated, which should not have happened. The author added more layers to the CNN to address this issue, implemented data augmentation, and then trained the model again.

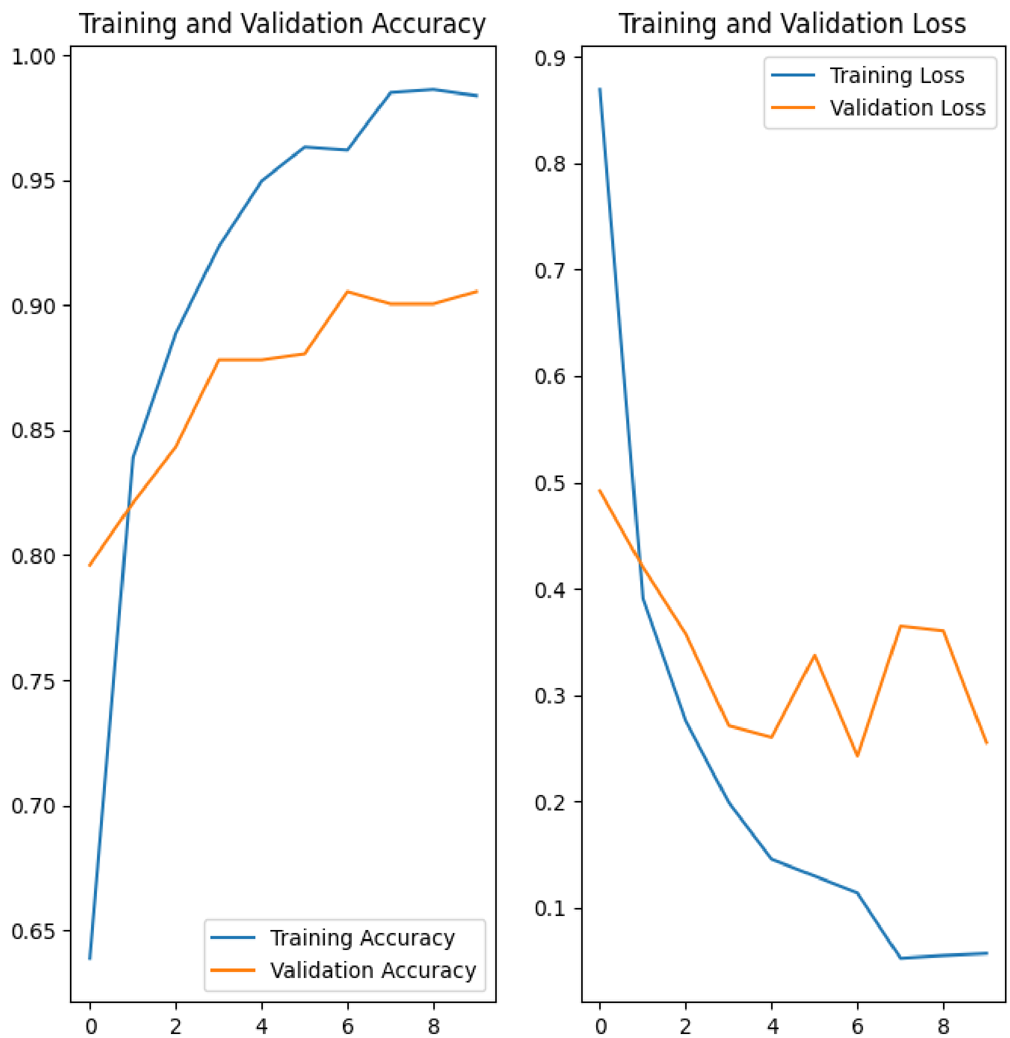

Figure 5. Results of the new image classification model, showing training loss, training accuracy, validation loss, and validation accuracy

After training for 20 epochs, the loss was 0.0537, the accuracy was 0.9838, the validation loss was 0.2554, and the validation accuracy was 0.905. Figure 5 shows how the model's performance improves over each training epoch. The training loss decreases, and the accuracy increases over the epochs while ensuring that the validation loss and accuracy also improve. At some points, it remains stable to avoid overfitting. Overall, this model improves both training and validation metrics.

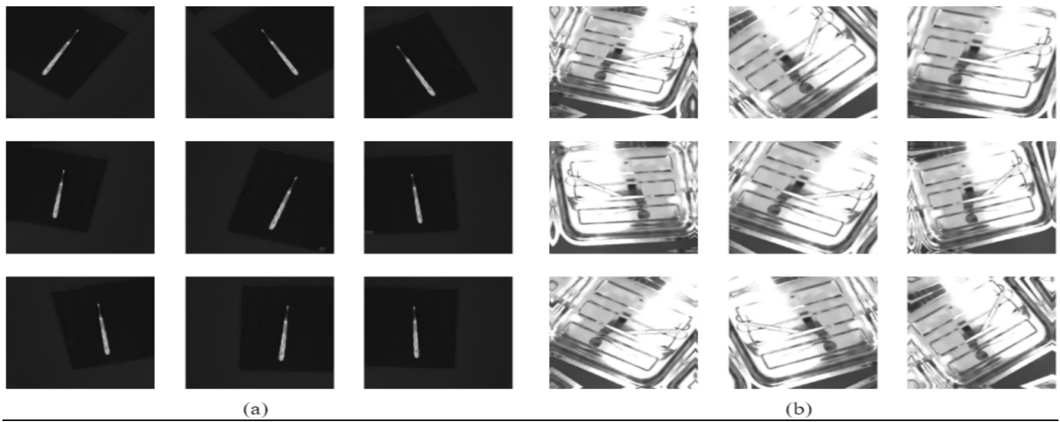

Figure 6. (a) Example of Dropout being implemented on an image of a scalpel. (b) Example of Dropout being implemented on an image of a curved mayo scissor

Dropout was utilized to improve the model further to prevent overfitting even further as the model learns over multiple epochs, a regularization technique commonly used in deep learning models. Dropout helps ensure the model learns to generalize well to unseen data by reducing the risk of memorizing noise or irrelevant patterns in the training data. During training, randomly selected neurons are temporarily removed or “dropped out” of the network with a certain probability (typically between 0.2 and 0.5). This means that these neurons do not contribute to the forward pass of the network and are excluded from the calculation of gradients during backpropagation. As a result, the network becomes less sensitive to the specific weights of individual neurons. It encourages the network to learn more diverse and independent representations of the input data, leading to better generalization performance on unseen data.

Final Model Accuracy

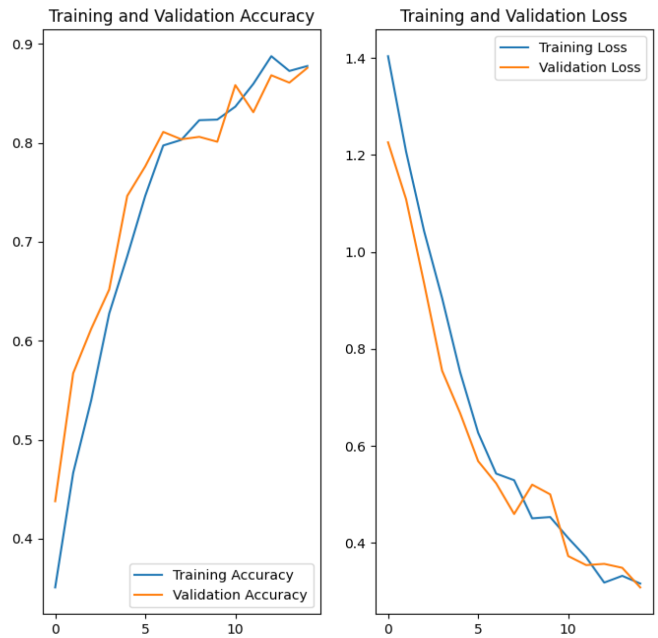

Figure 7. Results of the final image classification model, showing training loss, training accuracy, validation loss, and validation accuracy

As shown in Figure 7, the loss, both of the training and validation sets, decreases over epochs, indicating that the model is improving in its ability to minimize errors. The accuracy (training and validation) increases over epochs, suggesting that the model is better at correctly classifying examples. Dropout likely contributed to preventing overfitting, as evidenced by the consistent improvement in validation accuracy alongside training accuracy. The training time per epoch appears relatively stable despite the addition of dropout, which is expected as dropout only affects the forward pass during training.

Results

In the initial training phase, the model demonstrated promising results, achieving a final loss of 0.0089 and an accuracy of 0.9979 after 20 epochs. However, despite high accuracy on the training set, the model struggled to generalize to unseen data, as indicated by a validation loss of 0.3730 and a validation accuracy of 0.9375. To verify that the model failed to generalize well to unseen data, the model was trained again. The subsequent training phase showed similar results, with the loss remaining low and the accuracy increasing to 100%. However, despite the high accuracy, the model's generalization capabilities remained suboptimal, as evidenced by a validation loss of 0.3148 and a validation accuracy of 0.9688. This discrepancy suggested overfitting, where the model memorized the training data rather than learning underlying patterns.

Additional layers were added to the CNN to address overfitting, and data augmentation techniques were implemented. This technique yielded significant improvements, demonstrating an improved generalization, with the validation loss being 0.2554 and a validation accuracy of 0.9055. To improve these statistics further, in the final training phase, Dropout regularization was introduced to mitigate overfitting. Dropout randomly excluded neurons during training, encouraging the model to learn more representations of the data. This regularization technique proved effective, resulting in a final validation loss of 0.3078 and a validation accuracy of 0.8756.

Discussion

The integration of CNNs into surgical procedures represents a significant advancement in ensuring patient safety by improving surgical instrument count accuracy. The results obtained from the experiments conducted in this study demonstrate the efficacy of machine learning algorithms, particularly CNNs, in addressing the challenge of RSIs. One of the key findings of this study is the ability of the CNN-based model to achieve high accuracy in surgical tool recognition, even when trained on limited data scenarios. Despite the challenges posed by data scarcity and the complexity of surgical environments, this model demonstrates robust performance, as shown by its high accuracy and precision metrics. Through meticulous preprocessing, careful selection of hyperparameters, and the incorporation of advanced techniques such as data augmentation and Dropout regularization, the author was able to mitigate issues such as overfitting and improve the model's generalization capabilities. Moreover, the proposed approach addresses the immediate need to improve surgical instrument count accuracy and opens up possibilities for future advancements in the field. By utilizing surgical simulators and synthetic data generation techniques, the diversity and size of the training datasets can be expanded, further enhancing the model's performance and adaptability to diverse surgical scenarios. Additionally, ongoing technological advancements, such as integrating AI-powered cameras and real-time monitoring systems into surgical procedures, offer promising opportunities for enhancing patient safety and reducing the risk of RSIs.

Looking ahead, the successful implementation of machine learning algorithms in surgical tool recognition has the potential to revolutionize surgical practice. With continued research and development efforts, including interdisciplinary collaborations between computer scientists, medical professionals, and ethicists, existing challenges can be overcome and pave the way for the widespread adoption of CNN-based models in real-world surgical settings. Furthermore, as computer-aided surgery continues to evolve, there is hope for developing more sophisticated and intelligent systems that assist surgeons in tool recognition and enhance overall procedural efficiency and patient outcomes.

Conclusion

The research emphasizes the transformative potential of CNNs in addressing the critical issue of RSIs, an often overlooked problem. By synthesizing insights into how Machine Learning could potentially revolutionize RSIs, throughout the paper, the author has elucidated the challenges associated with manual counting methods and the potential of CNNs in automating surgical tool recognition. Exploring existing prevention strategies, including standardized protocols and technological innovations, has provided valuable context for understanding the urgency of improving current practices. This paper's empirical findings have yielded promising results, utilizing a CNN-based model trained on a labeled surgical tools dataset. Despite the inherent challenges of data scarcity and model generalization, this approach has demonstrated robust performance in accurately identifying and classifying surgical instruments. By leveraging advanced techniques such as data augmentation and dropout regularization, the paper has mitigated problems such as overfitting and enhanced the model's adaptability to real-world surgical scenarios.

Because of this study, ways exist to prevent retained surgical tools and enhance patient safety by implementing advanced machine learning algorithms. A model has been developed to characterize these tools, and the next step is to combine this model with hardware components to be used in real-time scenarios. In the future, there is immense potential for further research and innovation in this field. This study lays the groundwork for future endeavors to prevent RSIs through the widespread adoption of machine learning technologies. Using surgical simulators and synthetic data generation techniques could expand the diversity and scale of training datasets, thereby improving the efficacy and reliability of CNN-based models. For example, while training, surgeons can focus on more critical aspects of the procedure, knowing that instrument counts are being accurately monitored. Furthermore, the methodologies and insights gained from this research could be potentially applied to other areas within healthcare where accurate monitoring and automation are essential. The proposed machine learning model would offer real-time monitoring of surgical instrument counts during procedures, which provides immediate feedback to the surgical team, allowing them to take corrective actions promptly and ensure a safer surgical environment. With this tool's aid, students can perform surgeries more accurately going forward. By automating instrument count tracking, surgical teams can streamline procedures, potentially reducing overall surgery time and improving efficiency. The seamless integration will potentially allow easy adoption without disrupting established workflows, making it a practical solution for surgical settings.

Limitations

The sample size was limited to only four surgical tools, which restricts the generalizability of the findings. The small number of tools analyzed means that the results may not represent the broader range of surgical instruments used in practice.

Acknowledgments

I would like to thank Brian Zhu, Yocheved Kopel, Varun Madabushi, Jessica Goodard, Leila Chawkat, Nadia Hussain, and Raheel Haroon for their guidance and support throughout this project.

References

Bamba, Y., Ogawa, S., Michio Itabashi, Shingo Kameoka, Okamoto, T., & Yamamoto, M. (2021). Automated recognition of objects and types of forceps in surgical images using deep learning. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-01911-1

Bamba, Y., Ogawa, S., Michio Itabashi, Shindo, H., Shingo Kameoka, Okamoto, T., & Yamamoto, M. (2021). Object and anatomical feature recognition in surgical video images based on a convolutional neural network. International Journal of Computer Assisted Radiology and Surgery, 16(11), 2045–2054. https://doi.org/10.1007/s11548-021-02434-w

Chiew, C. J., Liu, N., Ting Hway Wong, Yilin Eileen Sim, & Hairil Rizal Abdullah. (2019). Utilizing Machine Learning Methods for Preoperative Prediction of Postsurgical Mortality and Intensive Care Unit Admission. Annals of Surgery, 272(6), 1133–1139. https://doi.org/10.1097/sla.0000000000003297

Chinedu Innocent Nwoye, Yu, T., Cristians González, Seeliger, B., Pietro Mascagni, Mutter, D., Marescaux, J., & Padoy, N. (2022). Rendezvous: Attention mechanisms for the recognition of surgical action triplets in endoscopic videos. Medical Image Analysis, 78, 102433–102433. https://doi.org/10.1016/j.media.2022.102433

García-Peraza-Herrera, L. C., Fidon, L., D'Ettorre, C., Stoyanov, D., Vercauteren, T., & Sébastien Ourselin. (2021). Image Compositing for Segmentation of Surgical Tools Without Manual Annotations. IEEE Transactions on Medical Imaging, 40(5), 1450–1460. https://doi.org/10.1109/tmi.2021.3057884

Gawande, A. A., Studdert, D. M., E. John Orav, Brennan, T. A., & Zinner, M. J. (2003). Risk Factors for Retained Instruments and Sponges after Surgery. *The New England Journal of Medicine*, *348*(3), 229–235. https://doi.org/10.1056/nejmsa021721

Jaafar Jaafari, Douzi, S., Khadija Douzi, & Badr Hssina. (2021). Towards more efficient CNN-based surgical tools classification using transfer learning. Journal of Big Data, 8(1). https://doi.org/10.1186/s40537-021-00509-8

Lam, K., Chen, J., Wang, Z., Iqbal, F., Darzi, A., Lo, B., Sanjay Purkayastha, & Kinross, J. (2022). Machine learning for technical skill assessment in surgery: a systematic review. Npj Digital Medicine, 5(1). https://doi.org/10.1038/s41746-022-00566-0

Lehr, J., Kelterborn, K., Briese, C., Schlueter, M. A., Kroeger, O., & Krueger, J. (2023). Image-based recognition of surgical instruments by means of convolutional neural networks. International Journal of Computer Assisted Radiology and Surgery, 18(11), 2043–2049. https://doi.org/10.1007/s11548-023-02885-3

Lee, J.-D., Chien, J.-C., Hsu, Y.-T., & Wu, C. (2021). Automatic Surgical Instrument Recognition—A Case of Comparison Study between the Faster R-CNN, Mask R-CNN, and Single-Shot Multi-Box Detectors. Applied Sciences, 11(17), 8097–8097. https://doi.org/10.3390/app11178097

Liu, K., Zhao, Z., Shi, P., Li, F., & He, S. (2022). Real-time surgical tool detection in computer-aided surgery based on enhanced feature-fusion convolutional neural network. Journal of Computational Design and Engineering, 9(3), 1123–1134. https://doi.org/10.1093/jcde/qwac049

Marzullo, A., Moccia, S., Catellani, M., Calimeri, F., & Elena De Momi. (2021). Towards realistic laparoscopic image generation using image-domain translation. Computer Methods and Programs in Biomedicine, 200, 105834–105834. https://doi.org/10.1016/j.cmpb.2020.105834

Praveen SR Konduri, & Siva, G. (2024). Full resolution convolutional neural network based organ and surgical instrument classification on laparoscopic image data. Biomedical Signal Processing and Control, 87, 105533–105533. https://doi.org/10.1016/j.bspc.2023.105533

Prellberg, J., & Kramer, O. (n.d.). Multi-label Classification of Surgical Tools with Convolutional Neural Networks. Retrieved December 11, 2023, from https://arxiv.org/pdf/1805.05760.pdf

Recommended Standard of Practice for Counts. (n.d.). https://www.ast.org/uploadedFiles/Main_Site/Content/About_Us/Standard%20Counts.pdf

Tracking surgical instruments with AI: A new approach to patient safety. (2023, October 25). Healthcare IT News. https://www.healthcareitnews.com/news/tracking-surgical-instruments-ai-new-approach-patient-safety

Weprin, S., Fabio Crocerossa, Meyer, D., Maddra, K., Valancy, D., Osardu, R., Hae Sung Kang, Moore, R. H., Carbonara, U., Kim, F. J., & Autorino, R. (2021). Risk factors and preventive strategies for unintentionally retained surgical sharps: a systematic review. *Patient Safety in Surgery*, *15*(1). https://doi.org/10.1186/s13037-021-00297-3

Zejnullahu, V. A., Bicaj, B. X., Zejnullahu, V. A., & Hamza, A. R. (2017). Retained Surgical Foreign Bodies after Surgery. *Open Access Macedonian Journal of Medical Sciences*, *5*(1), 97–100. https://doi.org/10.3889/oamjms.2017.005