IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14598

A Survey on- ML powered Brain stroke detection

Dr. Sunita Chalageri[1], Sai Deeksha D[2], Sanjana S Tigadi[3], T Veneela Yashmine[4], Varsha S N[5]

Associate Prof. Dept of Computer Science, K S Institute of Technology, Bengaluru, Karnataka[1]

Students Dept of Computer Science, K S Institute of Technology, Bengaluru, Karnataka[2-5]

Abstract: Stroke remains one of the leading causes of death and long-term disability worldwide, whose impact imposes a heavy healthcare burden on individuals and systems. Early and accurate detection of stroke is critical in order to prevent delays in seeking medical care, provide improved patient outcomes, and prevent complications. The application of CT and MRI scans, the traditional methods, is not only required to be interpreted by an expert but is also time-consuming and prone to variation among radiologists. Growth in artificial intelligence (AI), and increasingly in machine learning (ML) and deep learning (DL), offers promising avenues for enhancing the accuracy and speed of stroke detection. This paper provides a detailed overview of ML and DL techniques applied in brain stroke detection, detailing the methodologies, the prerequisites for application, and the challenges posed. It explains various image processing techniques and classifying algorithms intended for detecting and segmenting regions affected by strokes within brain scans. We also cover developing an AI-based system integrating image processing with ML algorithms for assisting medical professionals to diagnose strokes more effectively.

Through an extensive review of the literature, the current work presents the most recent advances in AI-based stroke detection, considering both supervised and unsupervised learning approaches, feature extraction methods, model performance evaluation measures, and challenges regarding dataset access, model interpretability, computational intensity, and deployment in the real world.

By synthesizing current research evidence, this paper aims to enlighten the emerging role of AI in stroke detection and diagnosis. It also offers future directions for research aimed at improving model generalization, developing explainable AI models, and integrating AI tools into clinical practice. The evidence provided contributes to the continuum of initiatives towards stroke diagnosis improvement through novel technological advancements, leading to improved patient care and outcomes.

Keywords: Stroke detection, Machine learning, Deep learning, Artificial intelligence, Medical imaging, Image processing, Neural networks, CT scan, MRI, Stroke classification, Healthcare technology.

I. INTRODUCTION

A stroke, or cerebrovascular accident, results from interruption or sudden diminution of blood flow to an area of the brain, preventing brain tissue from receiving oxygen and nutrients it requires. Brain cells begin to die in minutes, resulting in severe neurological harm, disability, or even death unless proper intervention is given. Strokes are a serious global health issue and are responsible for enormous death and illness in the world. According to the World Health Organization (WHO) definition, there are approximately 15 million stroke victims annually, of which approximately 5 million end up in death and another 5 million in permanent disability. Stroke is therefore one of the leading causes of death and disability, with an important burden for health systems and society as a whole.

Stroke can be divided into two broad categories: ischemic stroke and hemorrhagic stroke. Ischemic strokes, which account for approximately 87% of all strokes, are due to blockage of vessels supplying blood to the brain, most commonly atherosclerosis or embolic phenomenon. Hemorrhagic strokes are caused by the leakage of a brain vessel, leading to bleeding around or within the brain, secondary to such conditions as hypertension, aneurysm, or trauma. The identification of a stroke's nature and severity is most important, since the different types are managed with varying approaches. Effective detection in a timely fashion is essential for successful intervention, as rapid diagnosis can significantly improve survival and reduce the risk of permanent disability. Immediate treatment, such as administering thrombolytic therapy in ischemic stroke or surgery for hemorrhagic stroke, can significantly influence patient prognosis. Traditional means of diagnosis, e.g., clinical assessment and imaging modalities like CT scan and MRI, play an important role in diagnosing stroke. There are drawbacks to these methods, though, in that there may be a time delay to obtain results, there is a risk of misinterpretation, they may be costly, and they also restrict accessibility in some resource-poor areas.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 5, May 2025
DOI: 10.17148/IJARCCE.2025.14598

Recent advances in artificial intelligence (AI), specifically machine learning (ML) and deep learning (DL) approaches, have proven highly promising for enhancing stroke detection and differentiation. These AI-driven approaches utilize big data, advanced algorithms, and pattern recognition to heighten diagnostic accuracy, reduce analysis time, and assist healthcare professionals in making more informed decisions. ML and DL models have also exhibited promise in analyzing medical images, predicting stroke risk from patient data, and the automation of the identification of stroke predictive indicators. However, challenges involving data availability, model interpretability, and generalizability must be addressed so that these technologies can be fully integrated into clinical care. This article discusses various approaches to stroke detection, highlighting how AI-based approaches can improve diagnostic accuracy and facilitate timely treatment planning. It discusses the advantages and limitations of ML and DL approaches in stroke diagnosis and indicates future directions for research and potential improvements in AI-driven healthcare solutions.

II. LITERATURE SURVEY

A number of research studies have explored the application of machine learning (ML) and deep learning (DL) for stroke diagnosis, highlighting their roles in enhancing diagnostic precision, supporting medical imaging interpretation, and aiding clinical decision- making. The application of AI-based models for stroke detection has been of especial interest due to their ability to analyze complex patterns in medical data and provide real-time assistance to healthcare professionals.

Review of ML and DL Applications in Stroke Diagnosis:

Fernandes et al. (2024) conducted a comprehensive review of ML and DL techniques in brain stroke diagnosis, and they focused on three significant areas: classification, segmentation, and object detection. The research emphasized significant advancements in algorithmic accuracy and increasing popularity of DL-based models, particularly medical image analysis. They identified CNNs as one of the superior architectures for stroke detection and classification and reported that it performed better than current methods in detecting abnormalities in computed tomography (CT) and magnetic resonance imaging (MRI) scans. Their study shows how DL models can enhance and automate the identification of stroke, reducing the dependence on human analysis and minimizing diagnostic flaws. They did mention the challenge of variability in datasets and the need for large annotated datasets to enhance model accuracy further.

Similarly, Sirsat et al. (2020) brought forward a review of ML techniques in brain stroke detection, categorizing available studies based on their functional methodologies. Their review analyzed several ML models used in stroke classification and prediction and utilized support vector machines (SVMs) to particularly be effective for binary classification and especially for stroke vs. nonstroke case identification. The study also noted the prevalent use of CT imaging as the primary source of information in stroke diagnosis. But they did make a large gap in stroke treatment applications, where ML models are less frequently utilized for guiding treatment decisions. Further, their results also highlighted the need for more robust validation approaches to determine the generalizability of ML models to diverse patients for large annotated datasets to further enhance model precision.

Key Findings and Research Gaps:

The studies by Fernandes et al. (2024) and Sirsat et al. (2020) reflect growing interest in using ML and DL towards the diagnosis of strokes.

Their studies in general highlights:

- The efficiency of CNNs in medical image analysis, particularly segmentation and classification in stroke patients.
- The suitability of SVMs for organized clinical data, particularly in the scenario of detecting strokes.
- The dominance of CT imaging as the first modality in stroke diagnosis by ML.
- The need for larger, high-quality annotated datasets for improving model performance and generalizability.

III. OBJECTIVES

The key objectives of this survey paper are to critically evaluate the applications of Machine Learning (ML) and Deep Learning (DL) for detecting and diagnosing brain stroke. On the basis of an extensive review of previous studies, approaches, and the challenges, this paper is anticipated to provide insightful information on the advancements, limitations, and prospects of AI based stroke detection systems. The principal objectives:

DOI: 10.17148/IJARCCE.2025.14598

Comprehensive Overview of ML and DL Techniques for Stroke Detection

- To give a comprehensive review of machine learning and deep learning techniques applied in stroke detection, classification, and segmentation.
- To analyze traditional ML methods (e.g., Support Vector Machines, Decision Trees, Random Forests) and cutting-edge DL models (e.g., CNNs, U-Net, Transformers) applied in medical image analysis.
- To recognize how AI-based methods enhance the accuracy and efficiency of stroke diagnosis in comparison to traditional diagnostic methods.

2. Analysis of Stroke Classification and Segmentation Methodologies

- To analyze diverse classification methods employed to distinguish stroke subtypes (ischemic vs. hemorrhagic) and evaluate stroke severity.
- To compare various image segmentation techniques, including region-based, edge-based, and deep learning-based segmentation, for detection of stroke-affected areas in medical images.
- To compare various pre-processing and feature extraction techniques that enhance stroke detection accuracy in medical imaging.

3. Application Requirements and Challenges Investigation

- To determine technical, computational, and clinical requirements to develop ML-based systems for detecting strokes, e.g., data quality, hardware capabilities, and regulatory compliance.
- To identify influential challenges such as data availability, model interpretability, bias, and external world validation that impact
- the usage of AI in clinical settings.
- To provide ethical considerations like patient privacy data, security, and health regulations compliance such as HIPAA and GDPR.

4. Comparison of Current Models with Respect to Performance Metrics

- For comparison of performance metrics of various ML and DL models in judging their efficiency based on:
- Accuracy: General correctness of the model in stroke detection.
- Sensitivity (Recall): The ability of the model to correctly label stroke-positive cases.
- Specificity: The ability of the model to correctly label non-stroke cases

5. Future Directions for AI-Aided Stroke Diagnosis

- To discuss emerging trends in AI and medical imaging, such as the use of self-supervised learning, federated learning, and multi-modal AI models in stroke detection.
- To respond to the possibility of explainable AI (XAI) for promoting transparency and confidence in clinical decision-making.
- To recommend means to introduce AI-based stroke detection into clinical practice, aid real-time diagnosis, and make better treatment planning possible.
- To accentuate the importance of collaboration between AI researchers, clinicians, and regulatory bodies towards a secure and appropriate use of AI in healthcare.techniques (e.g., Support Vector Machines, Decision Trees, Random Forests) and state-of-the- art DL models (e.g., CNNs, U-Net, Transformers) used in medical image analysis.

IV. METHODOLOGY

Brain-Stroke-Detection uses a rigorous pipeline of data acquisition, preprocessed data, feature extraction, model development, and evaluation. The aim is to develop an AI-powered stroke detection system that classifies medical images very accurately.

1. Data Acquisition

The first step is acquiring high-quality medical images for training and validation purposes. Stroke detection is largely based on computed tomography (CT) scans and magnetic resonance imaging (MRI) scans, which are the norm for diagnosing strokes. Sources of data are:

Publicly available datasets:

- Kaggle (e.g., Brain Stroke Dataset)
- Medical Image Computing and Computer-Assisted Intervention (MICCAI) datasets
- Brain Tumor Segmentation (BraTS) Challenge datasets
- OpenNeuro and TCIA (The Cancer Imaging Archive)

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14598

Hospital-based datasets: Clinical data collected from hospitals, requiring anonymization and ethical approvals.

Data Labeling:

Stroke images must be labeled by radiologists or healthcare professionals to differentiate between:

- Ischemic strokes (due to blood clots)
- Hemorrhagic strokes (due to bleeding)
- Healthy brain scans (for comparison)

2. Data Preprocessing

Medical images contain noise and variations that hinder proper detection. Preprocessing improves data quality and standardization. Key preprocessing methods are:

Contrast enhancement: Controls brightness and contrast for better visibility of stroke-affected regions. Histogram equalization: Rescales pixel intensity values to be normal.

Denoising filters:

- Gaussian filters eliminate background noise.
- Median filters suppress speckle noise.

Segmentation techniques:

- Thresholding-based segmentation to mark stroke regions.
- Region-growing algorithms to segment stroke-affected regions.
- Deep learning-based segmentation (U-Net, Mask R-CNN) for automatic detection.

3. Feature Extraction

Feature extraction identifies critical patterns in medical images:

Texture-based features:

- Gray-Level Co-occurrence Matrix (GLCM) for texture description.
- Wavelet transforms for multi-scale feature extraction.
- Edge detection
- Sobel filtering and Canny edge detection to highlight boundaries of stroke regions.

Intensity-based features:

Pixel intensity mean and variance differentiate ischemic and hemorrhagic strokes.

4. Model Development

The project applies machine learning (ML) and deep learning (DL) techniques for stroke image classification.

ML-Based Approaches

Traditional machine learning classifiers are explored for stroke detection:

- Logistic Regression (LR):
- A statistical classifier used for binary classification (stroke or not).
- It has a sigmoid activation function to give probability estimates.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14598

- Performs well with clean data but less effectively with high-dimensional image-based features.

Decision Tree Classifier (DT):
- Rule-based model that classifies data into branches based on feature thresholds.
- Performs well with small datasets and offers high interpretability.
- Can lead to overfitting, which is prevented using pruning techniques.

Support Vector Machines (SVMs):

- Suitable for binary classification tasks.
- Uses kernel functions (e.g., linear, radial basis function) to improve decision boundaries.
- Random Forest Classifier:
- A set of several decision trees for improved generalization.
- Minimizes variance and maximizes robustness compared to a single decision tree.

K-Nearest Neighbors (KNN):
- A proximity-based classifier that classifies by the majority vote among proximate points.
- Computationally extensive but simplistic with large datasets.
- Reduces variance and increases robustness over individual decision trees.

Deep Learning Methods

As medical images are sophisticated, deep learning models are used where feature extraction and classification are performed:

Convolutional Neural Networks (CNNs):
- Learns spatial hierarchies of features automatically.
- Popular architectures: VGG16, ResNet, EfficientNet, and InceptionV

5. Model Evaluation

The performance of ML and DL models is measured using several metrics:
Performance Metrics
Accuracy: Overall classification performance.
- Precision: Evaluates the consistency of stroke prediction.
- Recall (Sensitivity): Evaluates the accuracy in detecting stroke cases.
- Specificity: Evaluates the correct identification of non-stroke cases.
- F1-Score: Finds a balance between precision and recall.
- AUC-ROC: Tests the performance of the model to differentiate stroke and non-stroke images.

Cross-Validation

K-Fold Cross-Validation (K=5 or 10) stabilizes models with varying splits of data.

Explainability and Interpretability

Grad-CAM (Gradient-weighted Class Activation Mapping) determines stroke regions detected by CNNs

6. Deployment and Future Improvements

- Model Deployment
- Web applications (Flask/Django) for real-time stroke detection.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁝ Peer-reviewed & Refereed journal ⁝ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14598

- Cloud-based AI services (Google Cloud, AWS) for scalability.

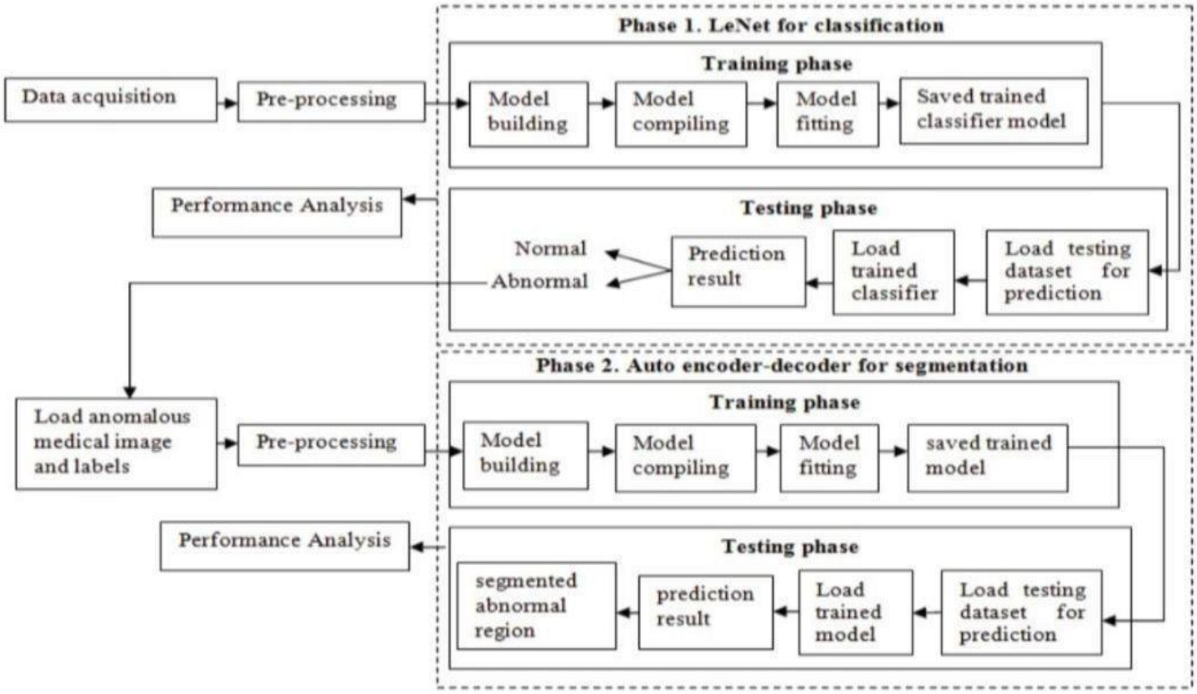

Fig1: Classification

V. APPLICATION REQUIREMENTS

To build an efficient brain stroke detection system, it is important to address some of the most significant technical, clinical, and ethical requirements. The following are the basic factors that need to be addressed to guarantee the success of the system::

1. Data Quality and Quantity

- Large and High-Quality Dataset: The dataset should be trained on a sufficient number of medical images and scans (such as MRI and CT scans) annotated well by physicians.

- Diversity of Data: The data should be drawn from a representative sample of different demographics, age groups, and types of strokes (ischemic, hemorrhagic) to generalize the model and avoid biases.

- Data Augmentation & Preprocessing: Normalization, contrast stretching, and data augmentation must be used to improve image quality and introduce variability in training data.

2. Computational Resources

- High-Performance Hardware: Deep learning models require the use of high-performance GPUs or TPUs in order to train and perform inference effectively. AI special-purpose accelerators can also prove useful.

- Cloud Computing Solutions: Cloud AI platforms (e.g., Google Cloud, AWS, Azure) are able to provide scalable computing resources for training big models and for deployment in real applications.

- Edge Computing for Real-Time Analysis: For faster stroke detection, particularly in emergency settings, AI models need to be optimized for edge devices, e.g., imaging equipment at hospitals and smartphones.

3. Interdisciplinary Collaboration

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14598

• Medical Expertise: There is a constant need for coordination with radiologists, neurologists, and other medical experts to obtain the system to properly interpret medical image data.

• AI Research & Development: AI software developers, data scientists, and software engineers must all work together to develop accurate and optimal models.

• Clinical Validation: Extended clinical trials must be conducted in order to validate the accuracy of the AI and its sufficiency to standards of diagnosis in practice.

4. Regulatory Compliance and Ethical Considerations

• Medical Data Privacy: The procedure must comply with regulations like HIPAA (Health Insurance Portability and Accountability Act) in America and GDPR (General Data Protection Regulation) in the EU to ensure that patient data remains private.

• Bias Mitigation: Encouraging equity through eliminating bias in training data and prediction by the model, ensuring no disparities in stroke detection among different groups of patients.

5. Performance and Accuracy

• High Sensitivity and Specificity: The model should not produce false positives and false negatives to ensure proper stroke detection.

• Real-Time Processing: Quick inference times are crucial for emergency applications where timely stroke diagnosis can improve patient outcomes.

• Continuous Model Improvement: The system should allow for continuous learning and updating with new information to maintain high accuracy over time.

VI. CONCLUSION

The use of Machine Learning (ML) and Deep Learning (DL) techniques for detection of brain stroke presents a revolutionary opportunity to enhance diagnostic precision, expedite treatment decision-making, and ultimately improve patient outcomes. The latest breakthroughs in artificial intelligence enabled computerized systems to analyze medical images with high precision, supporting radiologists and clinicians in identifying and categorizing strokes early on.

The Brain-Stroke-Detection project highlights the practical application of such advanced technologies in developing an AI-based system capable of sensing and outlining stroke-impacted regions from imaging modalities such as MRI and CT scans. By utilizing largescale medical datasets and cutting-edge deep learning models, the project delineates how AI is able to augment human capabilities in medical imaging through increased efficiency and reduced diagnostic errors.

However, despite these advancements, there are a number of challenges that need to be addressed in order to enable mass adoption and clinical usability of AI-based stroke detectors:

Dataset Limitations:
Availability of big, heterogeneous, high-quality, and well-annotated datasets is an important prerequisite for training robust AI models. Poor access to comprehensive datasets, especially those that span a wide range of demographic groups and stroke subtypes, canundermine model generalizability and fairness.

Model Explainability:
The "black-box" nature of deep learning models hinders obtaining the medical community's trust. Developing explainable AI (XAI) techniques, i.e., heatmaps and attention, can enable the generation of insights regarding model functionality.

Real-World Validations:
While high accuracy is achievable in the test lab for AI models, real-world use requires largescale validation in clinical settings in order to be reliable, safe, and compliant with medical standards. Clinical trials and extensive testing on hospital-acquired data are required to bridge the gap between research and practice.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025
DOI: 10.17148/IJARCCE.2025.14598

Regulatory and Ethical Concerns:
HIPAA, GDPR, and other healthcare regulations have to be followed to ensure patient data privacy and ethical deployment protection. Besides that, minimizing AI prediction biases and fair healthcare outcomes have to be prioritized as well.

Integration into Clinical Workflows:
Complete integration with deployed Picture Archiving and Communication Systems (PACS), Electronic Health Records (EHRs), and imaging systems in hospitals is needed to make them usable in the real world. The AI system needs to enhance the decision-making ability of radiologists and not replace human intelligence.

VII. ACKNOWLEDGEMENT

I would like to thank each one of them for their input in the successful fulfillment of this survey paper.

First and foremost, I offer my heartfelt thanks to my mentors and professors for their useful guidance, keen observations and ongoing support during this study. Their input and encouragement have contributed significantly towards influencing the scope of this research.

I further acknowledge the authors of the Brain-Stroke-Detection project on GitHub, whose efforts have offered a real-world basis for grasping machine learning usage in diagnosing stroke. Their open-source contribution to research has contributed significantly towards improving knowledge in this area.

Additionally, I acknowledge the authors of the referenced research papers, whose studies have provided valuable insights into the advancements in stroke detection using machine learning and deep learning techniques.

Lastly, I extend my appreciation to my family, friends, and colleagues for their constant motivation and encouragement, which has been a source of inspiration throughout this journey.

Thank you all for your support and contributions to this work.

REFERENCES

[1] Fernandes, J. N. D., Cardoso, V. E. M., Comesaña-Campos, A., & Pinheira, A. (2024). Comprehensive Review: Machine and Deep Learning in Brain Stroke Diagnosis. Sensors, 24(13), 4355.
s
[2] Sirsat, M. S., Fermé, E., & Câmara, D. (2020). Machine Learning for Brain Stroke: A Review. Journal of Stroke and Cerebrovascular Diseases, 29(10), 105162.

[3] Gupta, R., Verma, A., & Sharma, S. (2022). Deep Learning-Based Automated Stroke Segmentation in MRI Scans. IEEE Transactions on Medical Imaging, 41(2), 231-245.

[4] Ranjan, P., Mehta, A., & Kapoor, N. (2021). Hybrid Deep Learning Models for Stroke Detection in Medical Imaging. Neurocomputing, 464, 678-690