Chapter 7
Investigation on Generative Autoencoders for Hand Gesture Recognition

N. Bharanidharan
https://orcid.org/0000-0001-9064-8238
Vellore Institute of Technology, India

N. Mugunthan
https://orcid.org/0009-0008-6314-3128
Vellore Institute of Technology, India

B. Nitish
https://orcid.org/0009-0006-0558-4406
Vellore Institute of Technology, India

V. Puneeth
Vellore Institute of Technology, India

Kumar V. Vinoth
Vellore Institute of Technology, India

ABSTRACT

As a natural interaction method, hand gesture recognition has gained increased attention, particularly for human-computer interfaces. The goal of this research work is to develop a hand gesture detection system using machine learning techniques. Rather than making direct gesture predictions, this system compares the accuracy of several models in identifying hand gestures from dataset, thereby enabling real-time gesture-based interactions. To achieve this, we employed a combination of machine learning methods, neural networks, PCA for feature reduction and data pre-processing techniques. Algorithms like K-Nearest Neighbors, Decision Trees, SVM, and Autoencoders are used for classification. The dataset used in this work is the LeapGestRecog collection, which has 20,000 grayscale images of ten distinct hand gestures. Our best-performing model, SVM with an RBF kernel and auto-encoder, attained a remarkable accuracy of 99.98% on the test dataset, indicating its robustness in classifying hand gestures accurately.

DOI: 10.4018/979-8-3693-8332-2.ch007

1. INTRODUCTION

A cutting-edge technology called Hand Gesture Recognition allows machines to distinguish between human hand gestures, fostering smooth human-computer interaction. It has the potential to revolutionize several fields with its applications which includes like translating through Sign Language, Augmented Reality, Gaming as well as Robotic Control by capturing and analyzing the human hand gestures and its movements. This simple mode of communication bridges the gap between humans and machines, enhancing usability and accessibility across diverse fields (Kapitanov A et al., 2024). The range of Hand Gesture Recognition extends to a variety of fields, including healthcare, where it helps with rehabilitation therapy, and smart home systems that use gestures to operate electronic devices (Sharma & Singh, 2021). Hand gesture recognition has grown in significance as touchless interfaces have become more prevalent, particularly since the pandemic. By removing physical contact, it improves user experience while guaranteeing convenience and hygiene. Furthermore, the possibilities of gesture-based systems have been increased by hardware developments such as wearable sensors and depth cameras (Al-Hammadi et al., 2020). However, the challenge lies in accurately recognizing gestures amidst variations in lighting, hand shapes, and movements. In dynamic situations, robust systems need to be taught to recognize a variety of motions. To guarantee good performance even under real-world circumstances, this calls for the combination of potent algorithms and sophisticated computational models.

Machine learning is a crucial part of contemporary computational approaches because it enables systems to improve their performance by learning from data despite the requirement for written code (Chaudhary et al., 2013). Supervised learning, Unsupervised learning, and Reinforcement learning are the three primary classifications that machine learning falls under. In contrast to Unsupervised learning, which looks for hidden patterns in unlabeled data, Supervised learning uses labeled data for training. Reinforcement learning, on the other hand, is the process by which agents learn the best behaviors in an environment by trial and error. There are virtually countless uses for machine learning, starting with natural language processing and computer vision to predictive analytics as well as medical diagnostics. For example, recommendation systems use supervised learning to make product recommendations, whereas unsupervised learning uses clustering algorithms to assist categorize customers. Reinforcement learning has advanced autonomous driving and game AI, showcasing machine learning's ability to revolutionize a variety of industries. Because of its versatility and flexibility, machine learning is a perfect technique for hand gesture recognition, where the varieties like hand motions necessitates intelligent systems with generalization capabilities. Machine learning enables systems to dynamically adjust to novel gestures and users, opening the door for reliable, real-time implementations.

Deep Learning (DL), a part of Machine Learning that pushes boundaries regarding computational models by utilizing multiple layers of artificial neural networks. Deep Learning mimics the architecture of human brain to handle and evaluate complex patterns in data (Bobić et al., 2016). The types of Deep Learning consist of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), along with autoencoders, each tailored for specific tasks (Sharma S & Singh S, 2021). Deep Learning has found profound applications in fields like image recognition, speech processing, and medical diagnostics. Convolutional Neural Networks excel in image and video analysis, enabling breakthroughs in autonomous vehicles and medical imaging (Al Mudawi et al., 2024). Recurrent Neural Networks, with their ability to model sequential data, power advancements in language translation and speech synthesis (Sahoo et al., 2022). The flexibility of Deep Learning ensures its dominance in solving problems that were once deemed intractable (Al-Hammadi et al., 2020). The strength

Related Content

The Promise of Information and Communication Technology in Volunteer Administration
Kelly E. Proulx, Mark A. Hagerand Denise A. Wittstock (2018). *Technology Adoption and Social Issues: Concepts, Methodologies, Tools, and Applications (pp. 1388-1407).*
www.irma-international.org/chapter/the-promise-of-information-and-communication-technology-in-volunteer-administration/196734

Screen Time, Temporality, and (Dis)embodiment
Eduardo J. Santos, Ralph Ings Bannelland Camila De Paoli Leporace (2019). *Managing Screen Time in an Online Society (pp. 46-77).*
www.irma-international.org/chapter/screen-time-temporality-and-disembodiment/223053

Brain-Computer Interfaces in Neurorehabilitation: Advancing Cognitive, Motor, and Sensory Recovery Through Adaptive Neurotechnology
Jayashree Deka, G. Ashwin Prabhu, Majid Afsar Hussain, P. Tharaniya, Vikash Sharma, Mayuri Sanjay Mhaske, J. B. Kailashamand Sheron Joseph (2026). *Brain-Computer Interfaces for Neurorehabilitation (pp. 189-212).*
www.irma-international.org/chapter/brain-computer-interfaces-in-neurorehabilitation/409534

AI-Driven Helmet Compliance for Enhancing Human-Machine Collaboration in Smart Manufacturing: Smart Helmet and Traffic Violation Detection System
S. Mahalakshmi, K. Valarmathi, M. Meena, M. Shanmugapriyaand J. Thirumala (2026). *Navigating Human-Machine Collaboration in Smart Factories (pp. 25-64).*
www.irma-international.org/chapter/ai-driven-helmet-compliance-for-enhancing-human-machine-collaboration-in-smart-manufacturing/395090

Non-invasive vs. Invasive BCI for Clinical Applications
T. Gayathriand S. Kiruthika (2026). *Brain-Computer Interfaces for Neurorehabilitation (pp. 145-164).*
www.irma-international.org/chapter/non-invasive-vs-invasive-bci-for-clinical-applications/409532