Activity Monitoring and Unusual Activity Detection for Elderly Homes

Geocey Shejy, Manasvi Patwa, Bhavika Chattani, Sahil Deshmukh, Varnit Batheja

Vivekanand Education Society's Institute Of Technology, Mumbai, India

Abstract — The elderly population is growing rapidly worldwide, and with age comes an increased risk for falls, accidents, and other health issues. In order to ensure the safety and well-being of the elderly, it is crucial to have a system in place to monitor their activities and detect any unusual behavior. This paper discusses the importance of activity monitoring and unusual activity detection in elderly homes and suggests using current state-of-the-art technology for this purpose.

Keywords — activity monitoring, unusual activity detection, elderly care, ambient assisted living

I. INTRODUCTION

The elderly population is rapidly increasing in most countries and this trend is expected to continue in the future. The aging population is facing many challenges, one of which is falls, which are a common issue among the elderly. Approximately 30% of older adults fall at least once a year and these falls account for 75% of fall-related injuries. These concerns have led to the development of fall detection systems that can detect or prevent falls and provide quick and efficient help when they do occur. Despite these efforts, there are still relatively few commercially available fall detection systems on the market due to reliability issues, difficulty of installation and use, and concerns about privacy and cost.

Monitoring daily activities, or activities of daily living (ADLs), is crucial for detecting anomalies, promoting independent living, and diagnosing health problems. The development of remote health monitoring technology enables older adults to remain in their own homes rather than moving to expensive nursing homes or hospitals, promoting their independence. The proposed solution aims to provide a cost-effective and comprehensive healthcare option that allows elderly individuals to maintain their independence by remaining in their own homes.

The approach to detecting potential threats in smart homes is distinctive as it uses current activity knowledge to identify changes in usual patterns that may indicate the need for further investigation and response. This approach differs from conventional cyber-physical systems, which rely on a fixed set of parameters like time and location to determine the current context, which is the foundation for a variety of context-aware services, including security services. Many researchers adopt a strategy akin to the one outlined in this solution, searching for anomalies in sensor data patterns and interpreting those anomalies as threats to wellbeing.

Additionally, the solution includes an emergency button that can quickly notify emergency services or contacts through call or text in case of an emergency. This feature provides an added layer of security and peace of mind for older adults and their caregivers. In particular for the rapidly growing elderly population, the objective is to provide a cost-effective, discrete, and comprehensive healthcare solution with a minimal workforce for long-term health management and monitoring.

II. LITERATURE SURVEY

Ahatsham Hayat et al.[1] discussed that the recognition of elderly people's human activities has received relatively little research., despite the fact that many researchers have employed deep learning techniques to identify human activities. This study combines smartphone accelerometer and gyroscope data to follow the movements of senior persons in a variety of indoor and outdoor environments. The dataset includes routine activities like standing, lying down, walking, moving upstairs, and moving downstairs. Human activity recognition uses algorithms like k-Nearest Neighbors, Random Forest, Support Vector Machine, Artificial Neural Network, and Long Short-Term Memory Network. The training and testing datasets were subjected to two-fold and ten-fold cross-validation techniques to illustrate the effects of data modification. As compared to all other classification techniques, the recommended The Long Short-Term Memory Network provided an accuracy of 95.04 percent.

Gheorghe Sebestyen et al.[2] specified that the purpose of their paper is to list a number of routine activities performed by a person at home. and their research aims to look into the idea of creating a monitoring system that derives sensory information from several and various kinds of sources. They provide a method that includes three layers of data processing and recognition: the collecting and processing of raw data, the development of observations of actions, and the recognition of activity chains. For the purpose of activity chain recognition, a Hidden Markov chain approach is adapted. Their research focused more on the inference of different sensory inputs in order to identify complicated actions or activity chains. For this, they took into account sensors that pinpoint a person's location within the home, sensors attached to equipments, sensors that track a person's actions or states.

Eunju Kim et al.[4] states that activity recognition has the potential to greatly advance society, particularly in practical, human-centered fields like elder care and healthcare. This study concentrated on identifying basic human behaviors. The nature of human activities presents various problems, and the study of complex activities continues to be difficult and active. Understanding human activity includes both activity recognition and activity pattern discovery. The first is concerned with precisely detecting human activity using a predetermined activity model. In order to find activity patterns, an activity pattern discovery researcher first constructs an ubiquitous system and then analyzes the sensor data.

[5]An innovative technique for identifying human actions in health and social care services is suggested in the paper, which makes use of a depth camera and deep learning recurrent neural network (RNN). According to the authors, established approaches for identifying human activity in health and social care services are frequently constrained by a paucity of data and a requirement for pricey sensors.

The stages involved in preprocessing and feature extraction are described in the article, along with the process of gathering data with a depth camera. The authors next train and assess their model using a collection of films of human activities using a deep learning RNN. They claim to be quite accurate at diverse human behaviors including sitting, standing, and walking.
The authors come to the conclusion that their suggested approach has application potential in health and social care services since it can offer a non-intrusive and inexpensive way to identify human activity. Moreover, they suggest that their approach may be broadened to distinguish more complex tasks and track how activity patterns alter over time. Overall, the study makes the case that utilizing depth cameras and deep learning RNNs can enhance the ability to recognise human behavior in health and social care settings, offering a fresh method of observing and enhancing patient care.

[6]The system can transmit notifications to carers or emergency services in the case of an emergency. It is meant to detect trends and abnormalities in activities, such as movements, location, and vital signs. The authors claim that their system offers a non-intrusive and affordable way to monitor older people's health and well-being in their homes, which is something that is becoming more and more important as the world's population ages. The process of creating and testing the system, including the choice of suitable sensors and the creation of machine learning algorithms, is described by the authors. They claim to be quite accurate at identifying routine movements like sitting, standing, and walking as well as abnormalities like falls or extended periods of inactivity.

According to the article's conclusion, their technology has the potential to enhance geriatric care in smart home environments by offering a more complete and integrated solution. They propose that their method may be improved and combined with current smart home technology to offer a more practical and specialized approach to elder care.

III. RESEARCH METHODOLOGY

Our project includes following modules:

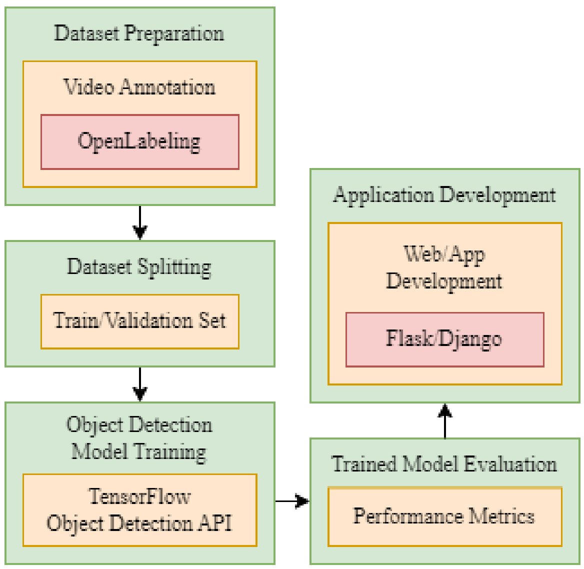

Fig. 1: Modular Diagram

Our methodology includes following steps:

1. Install TensorFlow and the TensorFlow Object Detection API: In this step, we installed TensorFlow and the TensorFlow Object Detection API, which is a collection of pre-trained object detection models and tools to train custom object detection models.

2. Prepare the dataset by annotating the RGB videos of elderly people performing usual activities. Label each frame of the video with the corresponding activity:
To prepare the dataset, we obtained a video dataset of elderly people performing various activities. We applied for the necessary license to download the video dataset. Then, we used OpenLabeling software along with different python libraries such as numpy, lxml, OpenCV, tqdm, etc. to annotate the RGB videos. OpenLabeling software was used to divide a video into a collection of frames, and then a bounding box was manually sketched around the concerned subject performing the activity. The tracker tracked the subject along with the bounding box which resulted in annotating all the frames. These annotated frames were then labeled with the corresponding activity.

3. Divide the dataset into training and validation sets:
After annotating the frames, we split the dataset into training and validation sets. This helps in evaluating the performance of the trained model on new data.

4. Train a specific object detection model using the TensorFlow Object Detection API, using the annotated dataset: To train a custom object detection model, we used the TensorFlow Object Detection API. We downloaded the pre-trained model and its corresponding config file from the TensorFlow Model Zoo and updated the config file with the necessary parameters such as the number of classes, the path to the training and validation datasets, and the batch size. Then, we trained the model using the annotated dataset. Once the training was complete, we saved the trained model to a file.

5. Create a Python script that uses the TensorFlow Object Detection API to perform object detection on a new video of an elderly person and classifies the activity being performed:

Finally, we created a Python script that uses the trained model and the TensorFlow Object Detection API to perform object detection on a new video of an elderly person and classify the activity being performed. The script takes a video as input and outputs the corresponding activity being performed in each frame of the video.

6. In summary, our methodology involved preparing the dataset by annotating the RGB videos of elderly people performing usual activities, dividing the dataset into training and validation sets, training a custom object detection model using the TensorFlow Object Detection API, and creating a Python script to classify activities in a new video using the trained model.

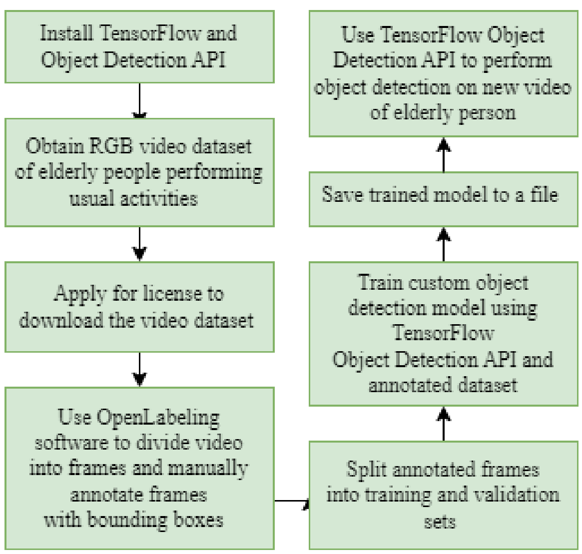

Fig. 2: Block Diagram

We have used 2 different models for activity detection:

1. TensorFlow SSD MobileNet V2 FPNLite 640x640 model
2. TensorFlow SSD MobileNet V2 FPNLite 320x320 model

Fig. 3 shows the comparison obtained by implementing these models using the same annotated video dataset.

Out of the two, a better accuracy was achieved by using the Tensorflow SSD MobileNet V2 FPNLite 320x320 model. This system is better than the existing ones as the existing systems use physical sensors which need to be attached to the body of a person. This may hinder their comfort, unlike our proposed system which will use cameras for real-time detection. Although, more accuracy can be achieved in our system by using complex models, from the TensorFlow model zoo, but they require high GPU and more memory which becomes a constraint for us. There is also scope for improvement in detecting small and medium objects.

	TensorFlow SSD MobileNet V2 FPNLite 640x640 model	TensorFlow SSD MobileNet V2 FPNLite 320x320 model
Speed	39	22
COCO mAP	28.2	22.2
Training steps	8000	10000
Average Precision at IoU = 0.50	75.1%	90.7%
Average Precision at IoU = 0.50:0.95 (overall)	33.7%	50.2%
Average Recall (how many objects did the model detect from the total objects in the dataset?)	54.7%	64.9%

Fig. 3: Comparison of 2 TensorFlow models used for Implementation

IV. CONSTRAINTS

1. The need for high-performance computing hardware is a significant constraint in our project, as the processing and analysis of large datasets require a significant amount of computational power. This can be a costly and time-consuming obstacle, as acquiring and maintaining high-performance hardware can be a significant investment.

2. Another constraint we face is the limited quantity of low-resolution datasets available for our analysis. These datasets are essential for developing and training our models, and a lack of sufficient data can lead to inaccurate or unreliable results.

3. Minimizing false positives is a major challenge we are facing. False positives occur when the model incorrectly identifies an activity which can result in false alarms and wasted resources. Finding ways to reduce false positives is essential for the reliability and effectiveness of our system.

V. RESULTS

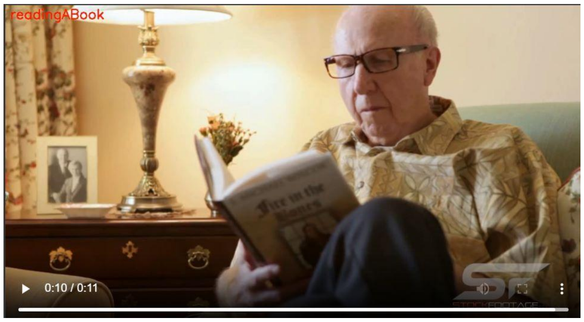

Fig. 4: Classification of the activity

Fig. 5: The above figure shows the identification of images and classification into respective action or activity categories.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.502
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.907
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.572
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.502
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.531
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.640
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.649
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.649
```

Fig. 6: Evaluation measures of the model

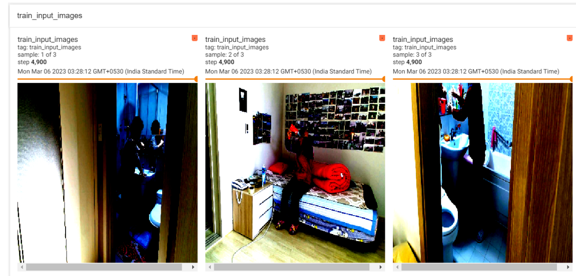

Fig. 7: Training images divided into 3 categories.

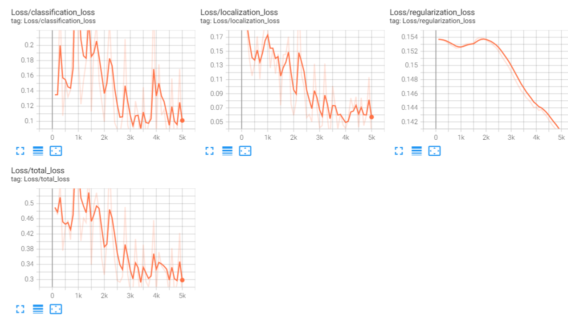

Fig. 8: Loss during training.

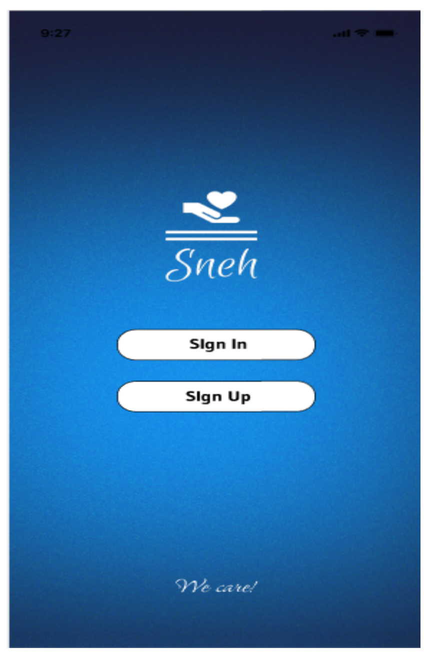

Fig. 9: App design

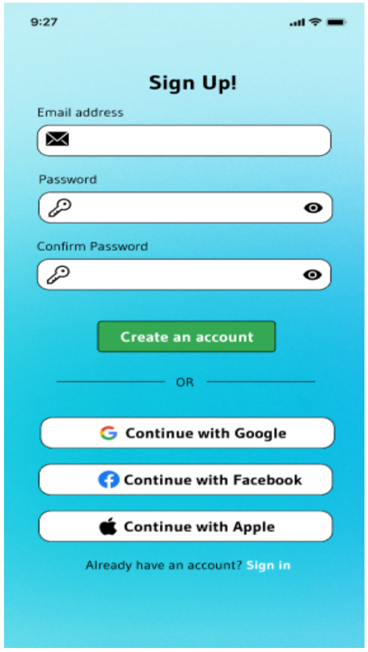

Fig. 10: Sign up page

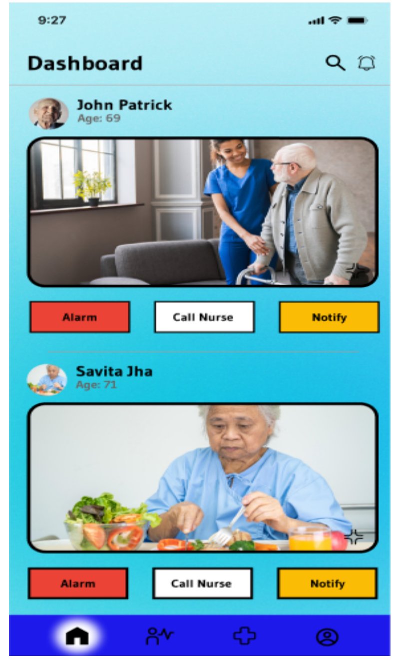

Fig. 12: Home page

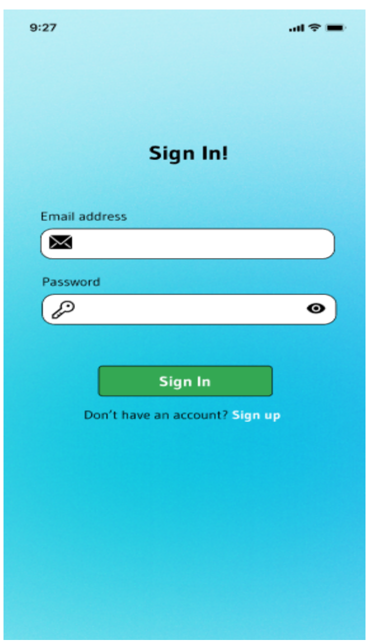

Fig. 11: Login page

Fig. 13: Activity page

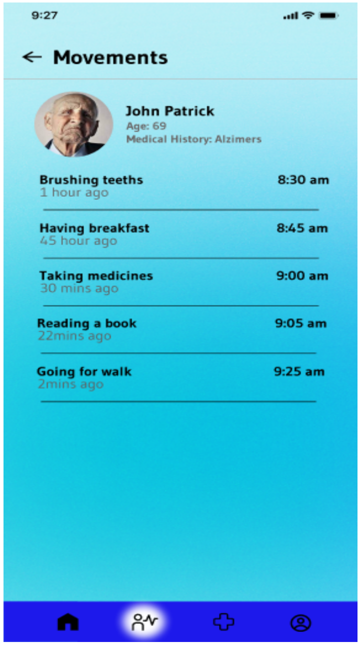

Fig. 14: Activity page

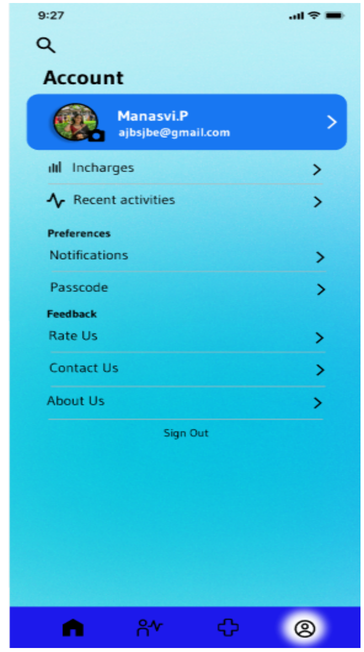

Fig.16: Profile of user

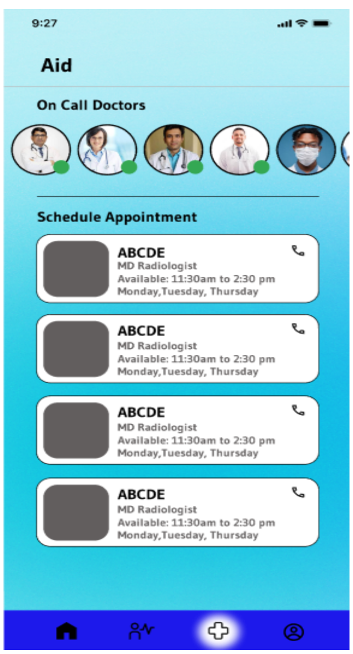

Fig. 15: Appointment page

VI. CONCLUSION AND FUTURE SCOPE

We used TensorFlow models for object detection. With the limited GPU, we could train the model with 2 architectures. The memory shortage of local storage constrained us to an accuracy of around 50%. The implementation of evaluation metrics and loss graphs through TensorBoard helped us better analyze our models. Our model performed moderately well on test images and on real-time detections via webcams. We are currently working on the development of the app from the design we have made. This will make it easier for caretakers and family members to monitor the elderly person's behavior and will help them receive real-time alerts for any unusual behavior.

A large fraction of the older population still cannot afford the formal paid care services provided by aged care facilities because of their high cost. The elderly may require regular, urgent medical assistance, which may be necessary to avoid fatal repercussions or even the person's death. To help seniors live independently, a detecting system is necessary. Since we'll be utilizing a camera to record an elderly person's activities, our system will be cost-effective, and every single movement of that individual will be documented. The cost of buying sensors like accelerometers, gyroscopes, etc. will be eliminated. Our system will have positive implications on the industry like improved safety, a better quality of life, improved care planning, etc.

The future scope of our project includes improvising the model's accuracy by increasing average precision for small and medium object size annotated images. Efficient storage will ensure using different architectures from the TensorFlow model zoo for better predictions.

VII. REFERENCES

[1] Ahatsham Hayat, Fernando Morgado-Dias, Bikram Pratim Bhuyan, and Ravi Tomar, *Human Activity Recognition for Elderly People Using Machine and Deep Learning Approaches, 2022.*
https://www.mdpi.com/2078-2489/13/6/275/pdf

[2] Gheorghe Sebestyen, Ionut Stoica, Anca Hangan, *Human Activity Recognition and Monitoring for Elderly People, 2016.*

https://sci-hub.hkvisa.net/10.1109/ICCP.2016.7737171

[3] Y. Charlon a b, W. Bourennane a b, F. Bettahar a b, E. Campo, *Activity monitoring system for elderly in a context of smart home, 2013.*
https://www.sciencedirect.com/science/article/abs/pii/S1959031812001509

[4] Eunju Kim; Sumi Helal; Diane Cook, *Human Activity Recognition and Pattern Discovery, 2009.*
https://ieeexplore.ieee.org/document/5370804

[5] S.U. Park, J.H. Park, M.A. Al-masni, M.A. Al-antari, Md.Z. Uddin, T.-S. Kim, *A Depth Camera-based Human Activity Recognition via Deep Learning Recurrent Neural Network for Health and Social Care Services, 2016*
https://www.sciencedirect.com/science/article/pii/S1877050916322943

[6] Nam S.B., Park S.U., Park J.H., Uddin MZ, Kim T.-S.
Accurate 3D human pose recognition via fusion of depth and motion sensors. International Journal of Future Computer and Communication. 2015
https://www.researchgate.net/publication/283661099_Accurate_3D_Human_Pose_Reco

gnition_via_Fusion_of_Depth_and_Motion_Sensors

[7] Jalal A, Uddin MZ, Kim T.-S.
Depth video-based human activity recognition system using translation and scaling invariant features for life logging at smart home. Consumer Electronics, IEEE Transactions on. 2012
https://www.researchgate.net/publication/326096546_Depth_Video-based_Human_Activity_Recognition_System_Using_Translation_and_Scaling_Invariant_Features_for_Life_Logging_at_Smart_Home

[8] Piyathilaka L, Kodagoda S.
Gaussian mixture based HMM for human daily activity recognition using 3D skeleton features.Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on. IEEE
https://ieeexplore.ieee.org/document/6566433

[9] Krizhevsky A, Sutskever I, Hinton GE.
Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems. 2012
https://www.researchgate.net/publication/267960550_ImageNet_Classification_with_Deep_Convolutional_Neural_Networks

[10] Duffner S, Berlemont S, Lefebvre G, Garcia C.
3D gesture classification with convolutional neural networks. Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. IEEE; 2014
https://ieeexplore.ieee.org/document/6854641