

LCAR-2021 Proceeding

*The Libyan Conference on Automation and
Robotics (LCAR-2021)*

No. 1, July 7-8, 2021 Tripoli/ Libya

المؤتمر الليبي للأتمتة والروبوتات

Libyan Conference on Automation and Robotics

Proceeding Contents

Conference Committees	3
Organizing Committee	3
Scientific Committee	3
Keynote Speakers.....	4
Evaluation of Automatic Method for the Detection and Quantification of the Abdominal Aortic Calcification Using Dual Energy X-ray Absorptiometry	5
Virtual Mouse Control through Natural User Interface	13
R-ORT: A Robust Object Recognition Technique from 2D Images Regardless the Viewing Angles	19
Speed Control of DC Motor Using Simulink With Arduino	25
Real-time Weather Station Monitoring System Through Wireless Sensor Network (WSN) Using off-the-shelf and Open-Source Components	31
Applying Multiple Deep Learning Models for Antipersonal Landmines Recognition	35
Design and Implementation of a Mini-Automated Teller Machine Replica	43
Implementation of the Computer Numerical Control Milling Machine	49
Modeling of Pneumatic Air Muscles for Rehabilitation Robotic Systems	55
How Crucial Is It for 6G Networks to Be Autonomous?	63

Conference Committees

Organizing Committee

- **Ali Ganoun**
University of Tripoli
- **Adel Korban**
University of Tripoli
- **Saleh Basha**
University of Tripoli
- **Dr. Abdusalam Moftah**
Faculty of Electrical &
Electronics Engineering, Beni
Waleed
- **Khaled Al-Ashouri**
Mared for Oil & Gas Services,
Tripoli
- **Azeddien Kinsheel**
University of Tripoli

Scientific Committee

- **Mohamed Elalem**
Elmergib University
maelalem@elmergib.edu.ly
- **Mohammed Elmusrati**
University of Vaasa
mohammed.elmusrati@uwasa.fi
- **Nuri Benbarka**
Universitat Tubingen, Germany
Nuri.benbarka@uni-tuebingen.de
- **Abdelhakim Deboucha**
Ecole Supérieure des Sciences Appliquées
d'Alger (ESSA-Alger)
a.deboucha@g.essa-alger.dz
- **Ausama Ahmed**
Elmergib University
ausama.ahmed@elmergib.edu.ly.com
- **Abdul Majeed Elbkosh**
Elmergib University
aalkbosh@elmergib.edu.ly
- **Ali Ganoun**
University of Tripoli
a.ganoun@uot.edu.ly
- **Alkhair Albarnawi**
College of Engineering
Technology-Houn
Alkhair_brno@yahoo.com
- **Abdulhamed Hwas**
University of Tripoli
ahamed.hawas@yahoo.com
- **Abdalla Elmanife**
College of Electrical &
Electronics Technology, Benghazi
a.almanfi@ceet.edu.ly
- **Abdullah Masrub**
Elmergib University
a.masrub@elmergib.edu.ly

Keynote Speakers

- ***Abdelmagid Hamouda***

Professor and Associate Dean for Academic Affairs

College of Engineering

Qatar University

Lecture Title: "Moving towards fifth industrial revolution - Rise of robots and drones"

- ***Sumeet S. Aphale***

Associate professor in Electrical Engineering, CEng, MIET, SMIEEE

School of Engineering

University of Aberdeen

Lecture Title: "Recent developments in control of nanopositioning systems"

- ***Othman Omran Khalifa***

Professor in Electrical and Electronics Engineering

International Islamic University Malaysia, IIUM,

Malaysia

Lecture Title: "Robotic vision object recognition using Deep learning"

- ***Ahmed Alkilani***

Associate Professors in Information Technology

University of Tripoli

- ***Ali Ganoun***

Associate Professors in Electrical and Electronics Engineering

University of Tripoli

Lecture Title: "Remote control and monitoring of oil fields"

- **السيد/ منجي نويرة**

مدير عام شركة سيا - تونس

Lecture Title: "تقنية الروبوت في الثورة الصناعية الرابعة"

Evaluation of Automatic Method for the Detection and Quantification of the Abdominal Aortic Calcification Using Dual Energy X-ray Absorptiometry

Karima M Elmasri

Nuclear Engineering Department, University of Tripoli, Tripoli, Libya
K.Elmasri@uot.edu.ly, Elmasrikm@yahoo.com

Keywords:

Abdominal aortic calcification;DXA; VFA, SVM

ABSTRACT

The aim of this is to evaluate an automatic method developed previously to detect and quantify the severity of the abdominal aortic calcification (AAC) by analysing the vertebral fracture assessment (VFA) images acquired using dual energy X-ray absorptiometry (DXA) in single energy mode.

The study Approval was obtained by Cardiff University, School of Engineering Ethics Committee. A set of 375 female VFA images of average age 74.17 ± 13.32 years, who underwent osteoporosis examination using DXA scanners between 2010 and 2017 were retrospectively collected. AAC severity was categorized into 4 classes (mild, moderate, severe and no-AAC). AAC was significantly correlated with the patient age; Pearson correlation was $r^2 = 0.82$, $p < 0.001$. The intraclass correlation (ICC) of two observers A and B to measure the AAC were 0.988 (95%, CI:0.985-0.991, $p < 0.001$ and 0.960 (95%, CI: 0.940-0.973, $p < 0.001$ respectively.

The agreement on four AAC categories on 100 VFA images between the readers was strong with $k = 0.634$. The best weighted accuracy, sensitivity and specificity were obtained using SVM classifier, which were 89.2%, 78.5% and 92.3% respectively for 4 AAC categories, and 88.6%, 86%, 90.4% for 3 AAC categories obtained by merging the moderate and mild calcification categories into one group.

I. Introduction

Atherosclerosis is a systemic disease described by the blood vessels walls thickening and hardening [1], [2]. Abdominal aortic calcification (AAC) has become a recognised predictor of cardiovascular diseases (CVD) during the last two decades and a manifestation of atherosclerosis [4]. In addition, a strong correlation between AAC and the coronary calcification has been shown in many studies [6]. AAC can be used to predict the myocardial infarction and stroke independently from the common clinical CVD risk factors [2], [7].

Several non-invasive diagnostic modalities were employed for AAC assessment, including radiography, computed tomography (CT) and ultrasound. Currently, the CT is the gold-standard technique for AAC quantification but it is limited by high radiation exposure [8], [9]. Dual energy X-ray absorptiometry (DXA) is the most common method generally used in medical imaging to detect osteoporosis through the bone mineral density (BMD) measurement and vertebral fractures. Vertebral fracture assessment (VFA) images acquired alongside BMD examination using DXA technology provide an inexpensive resource to diagnose the AAC.

VFA test has shown to be a plausible alternative for conventional radiography to detect the AAC because of its location just anterior to the spine site [10]–[12]. Schousboe et al. [11] have reported good agreement between lateral spine imaging with DXA and the conventional radiography in AAC detection.

The most popular semi-quantitative method produced by Framingham Heart Study [13] was used to assess AAC severity. This established and validated visual system (AAC-24) was used in many studies for AAC quantification in the lumbar region L1-L4 [5], [6], [11], [13]–[19]. In addition to the manual methods, several automatic approaches were proposed to detect AAC in radiographs [20], [21].

An automatic approach was proposed by the author of this paper [22] for AAC detection and quantification in VFA images obtained using two Hologic DXA scanners in single energy mode. The algorithm was applied and tested on 73 VFA images with clear evidence of calcification. Three calcification degrees of AAC: mild, moderate and severe were included in the selected VFA scans.

Considering that our former developed algorithm allowed segmentation and quantification of AAC automatically in images with three AAC categories, the author has hypothesised that the developed method would enable to assess the AAC on large number of VFA images including VFAs with non-calcified aorta.

II. MATREIALS AND METHODS

A. Study Population

This retrospective study was approved through the institutional review board of Cardiff, School of Engineering Ethics Committee. VFA images of 390 female patients referred for the investigation of osteoporosis between 2010 and 2017 at the

University Hospital of Wales, Cardiff, were defined from the Hospital archive and extracted from the Hologic software in DICOM format with sizes of (1104x402). Patient images were anonymised and each scan was given a specific number to avoid repetition.

The selected images exhibited a variety of calcification: starting with no calcification to severe calcification. Examples of VFA images of patients cropped for the abdominal region and edited for contrast enhancement are shown in Fig 1. These images in Fig 1 also scored by AAC-24. AAC severity was graded categorically by implementing a proposed method by Mohamed et al. [23].

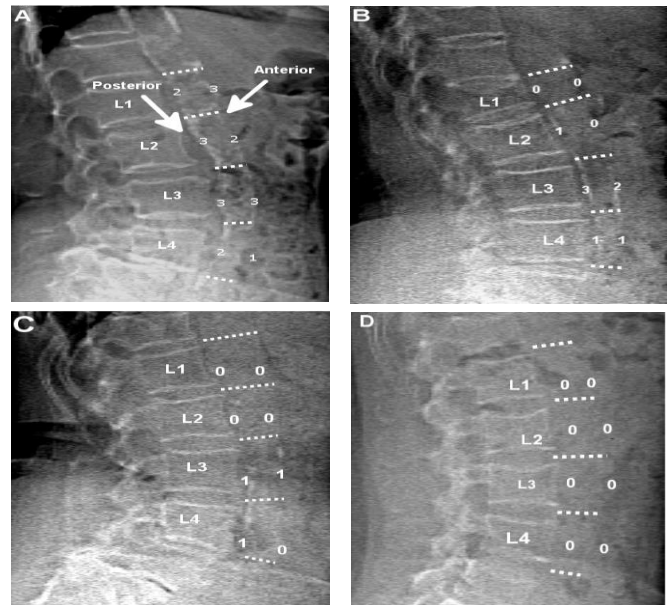


Figure 1. VFA images acquired by a DXA scanner; (A) a 91- year-old patient with severe AAC (AAC-24 =19), (B) a 74-year-old-patient with moderate AAC(AAC-24=8), (C) an 82 year-old- patient with mild AAC (AAC-24= 3), (D) a 38 year-old- patient with no AAC (AAC-24=0).

A set of 15 images with technical artefacts due to patient movements, significant anatomical artefacts and those with inadequate space anterior to the abdominal aorta for AAC scoring were excluded from this set. Fig 2 provides some examples of the excluded images and Table 1 lists the study data characteristics.

TABLE I. CHARACTERISTICS OF STUDY POPULATION

	Study Population n= 375	
	Average Mean	Standard deviation (SD)
Age, (years)	74.2	13.2
Height (cm)	155.0	7.2
Weight, (Kg)	61.1	11.5

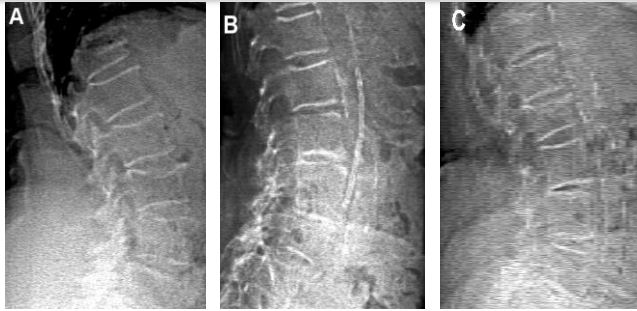


Figure 1. Excluded VFA images from the dataset due to;(A) inadequate space anterior to the abdominal aorta walls; (B) anatomical artefacts, (C) artefacts due the movement of the patient.

TABLE I Characteristics of study population

	Study Population n= 375	
	Average Mean	Standard deviation (SD)
Age, (years)	74.2	13.2
Height (cm)	155.0	7.2
Weight, (Kg)	61.1	11.5

B. Automatic Approach

The automatic approach of AAC assessment in VFA images presented in [22] was implemented in two stages. In the first stage, the abdominal aorta is automatically segmented using active appearance model (AAM) (34). In the second stage, calcification within the aortic walls is automatically quantified within the segmented area.

C. Automatic Segmentation of Aorta

A training set of 50 VFA images with clear evidence of calcification along each aorta walls were selected to train an AAM model for automatic segmentation. Creating an AAM requires manual placement of a set of landmarks at the points of interest in the training images. A set of 24 landmarks were placed in the corners and in the center of the end-plates of each four lumbar vertebrae (L1-L4). In addition, 32 landmarks were placed along the anterior and posterior walls of the calcified aorta adjacent to (L1-L4) as indicated in Fig 3. AAM was created on the basis of the landmark coordinates in the training set as well as the pixel intensities inside the shapes formed by the manually placed landmarks following the process described in [22].

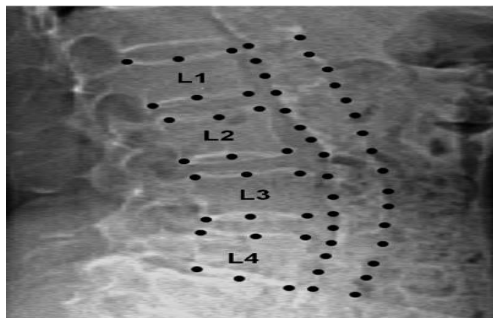


Figure 2 VFA image of a 91- year-old patient with clear calcified aorta labelled with 56 landmark points.

The automatic process of segmenting the aorta in the rest of the VFA images was commenced by initialising the shape of the lumbar segments and aorta to their average shape learnt during the AAM training process and placing the average shape approximately in the correct place on each of the remaining VFA images. The AAM search algorithm was then applied to each of the VFA images to find the best possible AAM fit as illustrated in Fig. 4.

The process of segmentation finishes by separating aorta from lumbar spine segments automatically as the largest segmented object [22]. Examples of detected aorta in the VFA images with AAC of different degrees are shown in Fig. 5.

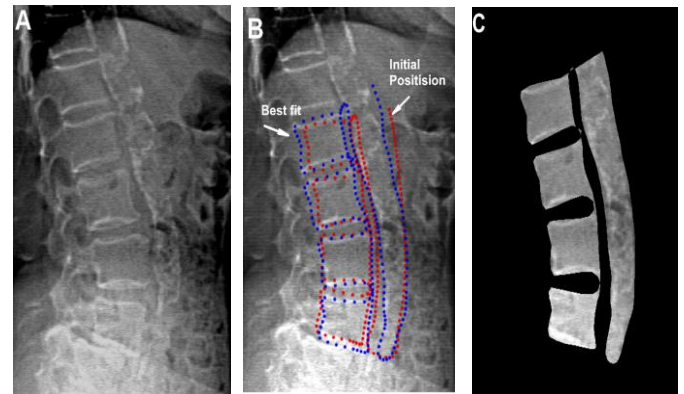


Figure 2. (A) VFA raw image; (B) optimising the best fit, initial position (red dots) the best fit (blue dots); (C) segmented lumbar vertebrae (L1-L4) and aorta region.

D. Automatic Quantification of AAC

Further image processing was required to minimize the variation of the image brightness and the effects of the quantity of soft tissue surrounding the aorta on the quality of the images, which makes application of a single threshold not suitable for separating calcified pixels from non-calcified.

Multilevel thresholding segments a grey level image into certain brightness parts separating the background and several objects [24]. The number of optimal thresholds was assigned by calculating the PSNR for each segmented aorta.

Every segmented image was split into eight levels by obtaining seven thresholds from multi-threshold. Thus, a new vector based on the maximum values was obtained and the maximum values in each quantization interval was assigned to the eight levels of the output image. Consequently, binary image was produced for each segmented aorta or its region of interest after this pre-processing stage [24].

Finally, connected regions within the binary image were automatically labelled, regions containing less than 50 pixels were considered as noise and removed as shown in Fig. 6.

Afterwards, quantitative features relevant to the AAC were extracted from the obtained segmented aorta.

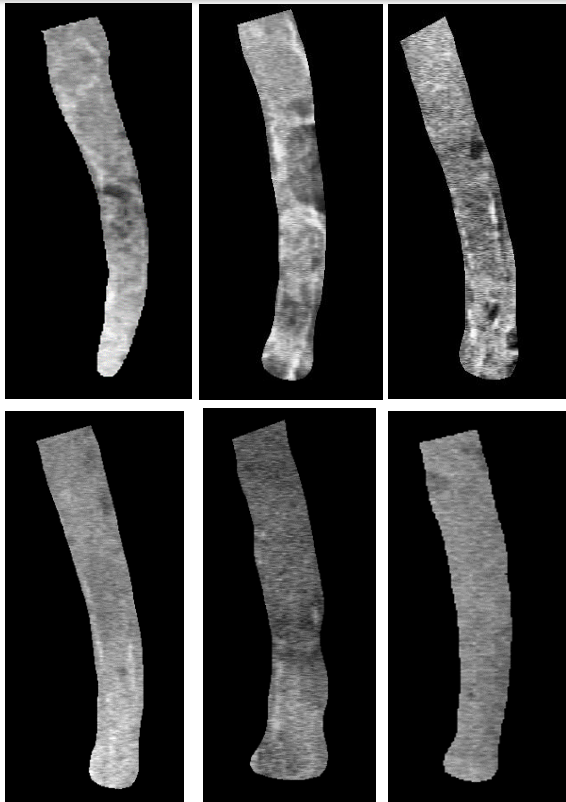


Figure 3. Various examples of segmented aorta with different degrees of AAC.

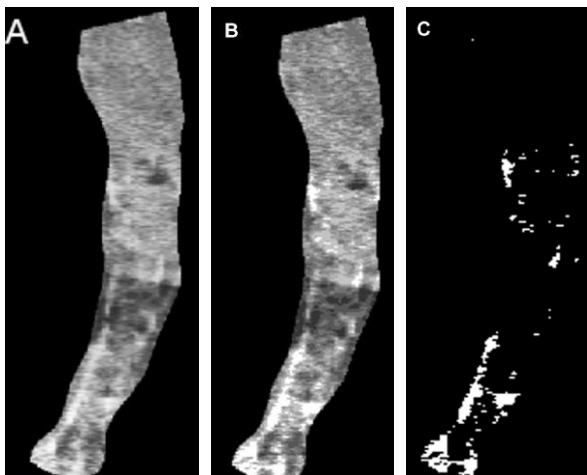


Figure 4. (A) Segmented aorta quantised using 7 threshold levels; (B), (C) binary images after thresholding.

The fraction of calcified pixels to the total number of pixels from every segmented aorta and seven other features obtained by calculation the grey-level co-occurrence matrix (GLCM) were extracted [25]. These features were: mean, variance, energy, entropy, homogeneity, contrast, and correlation.

E. Automatic Calcification Classification

1) Feature Selection

Informative features with the greatest discriminatory power were selected by applying feature selection algorithms to the extracted features. Feature selection usually does not change the data, so it is the best choice when data interpretation is essential, and this is the preferred choice for studying the segmented aorta features. Information gain and gini index algorithms [26] were used to analyse the features extracted from the segmented images [27]. This stage was used to enhance the performance of classifiers. The order of ranked features starting with the most important was as follows: the fraction of calcified pixels; pixels mean, entropy, energy, homogeneity, contrast, correlation, and variance respectively.

2) Automatic Classification

The performance of three popular classifiers, namely, Support vector machine (SVM) using linear kernel from LIBSVM, K-nearest neighbor (KNN) and random forest (RF) were tested. A K-folds cross validation method was conducted with 5 folds to partition the data set into two sets: training set to train the classifiers and test set to validate algorithm. Orange package software [28] was implemented in the classification stage. In this analysis a step of ranking of features were applied to evaluate the significance of the features.

Statistical measures such as accuracy, sensitivity, and specificity were used to describe each classifier performance. Sensitivity is the percentage of the positive cases which are correctly diagnosed, meaning the ability of the test to identify the category of the calcification (0, 1, 2, and 3). Specificity is the percentage of negative cases that are correctly classified as negative. AAC categories were tested one against the others. The difference between three classifiers also was investigated by implementing 5 folds 10 times for each classifier and unpaired t-test was conducted with 95% confident interval (CI) and $p < 0.05$ to detect significant statistical difference between the classifiers.

III. Statistical analysis

The statistical analysis was performed using Matlab (Mathworks, MA, USA) and SPSS (IBM, New York, USA). Calculated Pearson's correlation coefficient and 95% confident intervals (CI) were calculated to evaluate the relationship between patient age and AAC-24 scores, $p < 0.05$ was considered statistically significant [29].

The presence and severity of AAC were assessed for each VFA image by two observers. Reader A recorded AAC-24 scores twice for the whole VFA image set. Reader B scored the AAC-24 on 100 VFA images selected randomly and all AAC degrees were included. These readings were recorded using Matlab (Mathworks, MA, USA), applying contrast adjustment and digital zoom.

Intraclass correlation (ICC) between the two AAC-24 readings recorded was calculated.

Interclass variability of AAC scores recorded by the two readers was computed on set of 100 VFA images. To test the agreement

in evaluation of AAC categories between two readers, a simple agreement among readers was measured as a percent agreement and cohen's Kappa (κ) cross tabulations algorithm statistics were used [30], [31].

In general, a kappa coefficient of 0–0.20 is considered poor agreement, 0.21–0.4 is fair, 0.41–0.6 is moderate, 0.61–0.8 is strong, and >0.8 is considered about complete agreement [32].

In addition, the classification stage was conducted 10 times basing on k-folds. Statistical t-test also was conducted to test classifier's weighted average classification accuracy achieved by each classifier [32].

IV. Results

A. Data Analysis

The automatic segmentation method was successful in segmenting aorta in 369 out of 375 VFA images due to the presence of fracture in some vertebrae and poor spatial resolution. The distribution of grouped diagnoses in the AAC-24 system as follows: no-AAC was (127) and those who had severe AAC was the lowest number (64), the mild and moderate AAC instances were 93 and 85 respectively.

B. Correlation Between Patient Age and AAC

There was no calcification score recorded in patients aged < 40 years, in contrast, the majority of patients aged ≥ 40 years were given scores with some degree of calcification. The Pearson's correlation coefficient and 95% CIs between age and AAC-24 scores were, 0.82 (CI, 0.783 - 0.834, $p < 0.001$). Furthermore, about 56% of patients had AAC at level L4, this percentage was decreased at levels L3 to L1; 48%, 41% and 31%. Overall, the posterior aortic wall segments that was given scores from 1-3 was greater than the anterior aortic wall.

C. Validation of AAC-24 system

1) Intra-Observer Agreement of Calcification Scores

Table 2 summarises the two Readers intra-class correlation results.

	Number (n)	Intra-observer variability		P value
		AAC-24	4-AAC categories	
Reader A	369	0.988 (0.985 to 0.991)	0.986, (0.982 to 0.989)	<0.001
Reader B	100	0.960 (0.940 to 0.973)	0.953, (0.929 to 0.968)	<0.001

2) Inter-Observer Agreement of Calcification Scores

The agreement between Reader A and B on 100 VFA scores was measured for both AAC-24 points and AAC categories. There was a strong and significant correlation between the AAC-24 scores given by the two readers, $r^2=0.956$, $P<0.001$,

TABLE 1: INTER-OBSERVER OBTAINED FOR TWO READERS FOR AAC-24 AND AAC CATEGORIES WITH 95% CI.

Method	Reader A vs Reader B n=100			
	AAC-24	AAC category		
ICC	0.953,(0.9310.969)	0.934 (0.903 to 0.956)		
Kappa(κ)		0.634 (0.505 - 0.743)		
Agreement %		Class 0	Class 1	Class 2
		89%	73%	65%

D. Classification Accuracy and Feature Selection

1) Classification into 4-categories

Figures 8 shows the weighted average classification accuracy of the 4 AAC categories depending on the number of features used. The accuracy of SVM increased with the number of features to reach 89% using the first 6 features. RF accuracy decreased slightly after using 8 features, while k-nn classifier's accuracy decreased about 3% after using 8 features.

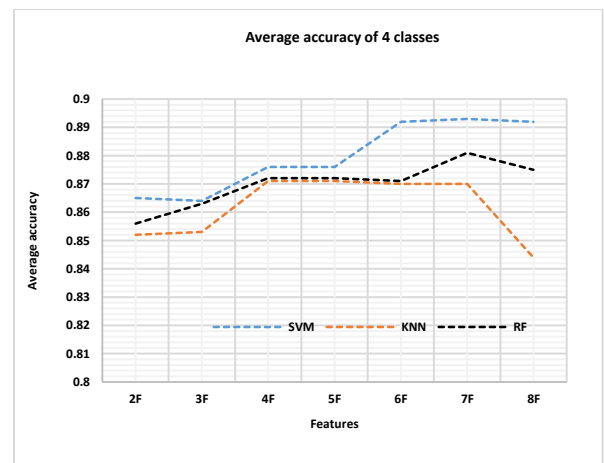


Figure 5. Average classification accuracy obtained with three classifiers with the number of features used for 4 AAC categories.

The best optimal weighted average accuracy, sensitivity, and specificity of each AAC category obtained by SVM, 0.91%, 0.86%, 0.83%. The classification was conducted by utilising the first 6 features with SVM, first 7 features with RF and KNN. The calculation was implemented 10 times with 5 folds; class 0 (no-AAC), class 1(mild-AAC), class 2 (moderate AAC) and class 3 (severe AAC) and the results averaged over all runs. The best accuracy, sensitivity and specificity were achieved by using SVM classifier.

There was a significant difference between the average accuracy achieved by SVM and KNN, 0.022(CI,0.016-.027, $p<0.001$); and between SVM and RF, 0.0116(CI,.004-0.189, $p=0.0033$); and between RF and KNN, .0102(CI,.0038-.0168, $p=.0043$).

2) Classification into 3-Categories

The two sets of images, corresponding to mild and moderate degree of calcification, were merged into a single set to enhance the sensitivity and specificity of classification. Mild and moderate groups of AAC-24 = (1-12) were merged as class 2 calcification. Classes 1 (no calcification, AAC-24 score of 0) and 3 (severe AAC, AAC-24 score > 12) remained the same. The classification results for three classes are shown in Fig 8. The best optimal weighted average accuracy, sensitivity, and specificity of every AAC category achieved by SVM, 0.92%, 0.86%, 0.96%. Classification was conducted by using 6 features for each classifier, 10 times with 5 folds.

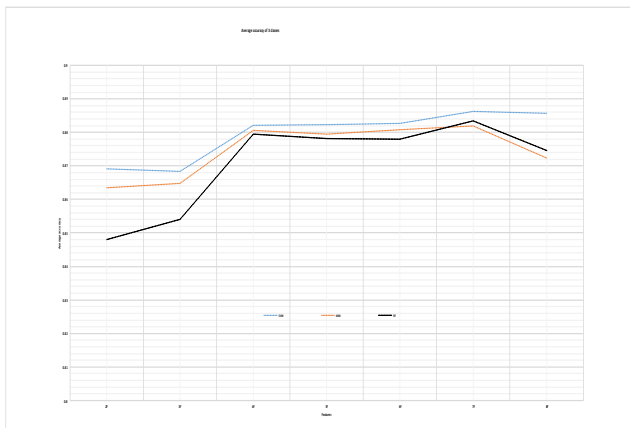


Figure 6. Average classification accuracy achieved by three classifiers with the number of features used for 3 AAC categories.

The three classifiers were tested with t-test and they revealed non-significant difference between three used classifiers: SVM and KNN, .0043(CI,-0.0014-0.010, $p=.128$); SVM and RF, 0.0028(CI,-0.0046-0.0102, $p=0.4403$; Knn and RF,0.0015(CI,-.005-0.0081, $p=0.63$).

V. Discussion

This study provides more evidence supporting DXA capability in evaluation of AAC through VFA images. A set of 375 VFA images obtained by using DXA scanners in single energy mode were retrieved from the Cardiff University Hospital archive. These images were obtained using DXA for osteoporosis and vertebral fractures diagnosis with low exposure to radiation. The study presents an automatic approach for segmentation of abdominal aorta and its surrounding area as well as quantification based on statistical analysis. A set of 15 images with technical artefacts and inadequate space anterior to the lumbar spine were excluded due to the difficulty to find the anterior and posterior of aorta within these images for segmentation purpose or for manual quantifying of AAC by using AAC-24 points.

To quantify AAC on VFA images, Active appearance model (AAM) was implemented for segmentation of the abdominal aortic walls and their surrounding areas. It has been proven in many previous works that AAM is a powerful algorithm for automatic medical images segmentation.

AAC-24 system points was strongly related to the age and the most common calcified deposits are detected in L4 segment as it was proven in many studies [33], [34]. The calcification was greater in the posterior wall than the anterior.

Overall, the AAM was robust and fast, the number of iterations required by the searching algorithm to complete this stage was less than 100 iterations as overall and it took only 23 seconds.

To quantify AAC, features extracted from segmented abdominal aortic region were used to measure the degree of calcification. Feature's selection algorithm was used in the classification stage to check the accuracy achieved by three classifiers SVM, RF and KNN.

The best accuracies obtained by the three classifiers SVM, KNN and RF were recorded for class 0, (91.4%, 92.2% and 91.1%) and class 3, (94.4%, 91.8%, 91.7%) respectively. Sensitivity was about 85% for no -AAC and severe AAC, and this was just under 70% for mild and moderate AAC calculated by SVM and less than 70% in KNN and RF. Specificity computed by three classifiers was >85% for all AAC classes.

Two classes, mild and moderate were merged and a new classification of AAC categories were conducted; these categories were no-AAC (AAC-24=0), mild and moderate (0 >AAC-24 <12) and high AAC (AAC-24 >12). The sensitivity and specificity of mild and moderate AAC were enhanced in all classifies.

The step of merging the mild and moderate calcification groups could improve the sensitivity and specificity with similar accuracies in the classifiers.

AAC-24 manual scoring system also was successfully applied based on many previous studies and the two readers have scored a set of 100 VFA in a consistent way, kappa correlation has shown a moderate agreement basing on four AAC categories. The algorithm could deal with the VFA noisy images and it is expected to be good for the images obtained by standard radiography as well.

As it was mentioned above that the algorithm was applied on a set of a small data set of 73 images good accuracy was achieved on this set. The accuracies obtained in this study were as good as those obtained for small set of images in our previous work and greater than 83% for all classes.

Clinically, the proposed method can be introduced into the medical imaging fields relevant to cardiovascular diseases and VFA images obtained for osteoporosis can be also used to predict vascular calcification basing on AAC.

This achieved results was good in this large set comparing with our previous work where the number was about 1/5 of current data set including images of patients with no AAC.

VI. Study Limitations

This proposed algorithm is limited with the calcification on the surrounding calcified areas which represents a calcification in some parts close to the aorta position. It is not applicable in VFA images which contain anatomical and some technical artefacts. The stages were not accurate as much as the high calcification stages where the sensitivity of low calcified aortic

walls was 60% AAC (class 1) obtained by SVM comparing with the severe AAC with sensitivity was $> 80\%$ for all classifiers.

VII. Conclusion and Future work

In conclusion, different degree of abdominal aortic calcification can be detected by DXA modality with single energy mode. The pattern of AAC distribution was similar to previously reported findings in many other studies, with most high calcified deposits were detected at L4 level decreasing towards L1.

The present approach may assist to identify patients with atherosclerosis before symptoms of cardiovascular development. It provides a fast and low-cost assessment of location and severity of calcification within the abdominal aorta.

In future studies, to enhance the outcome of the automatic approach, an interaction with the experts in this field can be conducted on the stage of land marking and AAC scoring to validate this algorithm.

References

- [1] J. Sanz and Z. a Fayad, "Imaging of atherosclerotic cardiovascular disease.," *Nature*, vol. 451, no. February, pp. 953–957, 2008.
- [2] R. Golestani *et al.*, "Abdominal aortic calcification detected by dual X-ray absorptiometry: A strong predictor for cardiovascular events.," *Ann. Med.*, vol. 42, no. 7, pp. 539–45, 2010.
- [3] L. Mosca *et al.*, "Evidence-based guidelines for cardiovascular disease prevention in women," *J. Am. Coll. Cardiol.*, vol. 43, no. 5, pp. 900–921, 2004.
- [4] P. W. Wilson *et al.*, "Abdominal aortic calcific deposits are an important predictor of vascular morbidity and mortality.," *Circulation*, vol. 103, no. Cvd, pp. 1529–1534, 2001.
- [5] C. R. Walsh *et al.*, "Abdominal aortic calcific deposits are associated with increased risk for congestive heart failure: The Framingham Heart Study," *Am. Heart J.*, vol. 144, no. 4, pp. 733–739, Oct. 2002.
- [6] M. Ganz *et al.*, "Distribution, Size, and Shape of Abdominal Aortic Calcified Deposits and Their Relationship to Mortality in Postmenopausal Women," *Int. J. Biomed. Imaging*, vol. 2012, pp. 1–8, 2012.
- [7] J. T. Schousboe, K. E. Wilson, and T. N. Hangartner, "Detection of aortic calcification during vertebral fracture assessment (VFA) compared to digital radiography.," *PLoS One*, vol. 2, no. 8, 2007.
- [8] R. W. Jayalath, S. H. Mangan, and J. Golledge, "Aortic Calcification," vol. 488, pp. 476–488, 2005.
- [9] M. Cecelja, M. L. Frost, T. D. Spector, P. Chowienczyk, and Received:, "Abdominal aortic calcification detection using dual-energy X-ray absorptiometry: validation study in healthy women compared to computed tomography.," *Calcif. Tissue Int.*, vol. 92, no. 6, 2013.
- [10] N. Schousboe, J.T., Vokes, T., Binkley, N., Genant, H.K., Sambrook, P. and Pocock, "Densitometric vertebral fracture assessment (VFA)," *Int. Osteoporos. Found.*, vol. 12, 1-24., 2010.
- [11] J. T. Schousboe, K. E. Wilson, and D. P. Kiel, "Detection of abdominal aortic calcification with lateral spine imaging using DXA.," *J. Clin. Densitom.*, vol. 9, no. 3, pp. 302–8, 2006.
- [12] P. N. John T. Schousboe, "Novel Uses of Dual Energy X - ray Absorptiometry (DXA): More than just bone density measurement John T . Schousboe , MD , PhD Park Nicollet Osteoporosis Center , Park Nicollet Clinic , Minneapolis , MN Division of Health Policy and Management , Universi," pp. 1–8.
- [13] P. W. F. W. Kauppila, Polak. Joseph F, L. Adrienne Cupples, Marian T. Hannan, Douglas P. Kiel, "New indices to classify location, severity and progression of calcific lesions in the abdominal aorta: a 25-year follow-up study.," *Atherosclerosis*, vol. 132, no. 2, pp. 245–50, Jul. 1997.
- [14] J. C. Witteman, F. J. Kok, J. L. van Saase, and H. a Valkenburg, "Aortic calcification as a predictor of cardiovascular mortality.," *Lancet*, vol. 2, no. 8516, pp. 1120–1122, 1986.
- [15] F. Bastos Gonçalves *et al.*, "Calcification of the abdominal aorta as an independent predictor of cardiovascular events: a meta-analysis.," *Heart*, vol. 98, no. 13, pp. 988–94, Jul. 2012.
- [16] C. J. O'Donnell *et al.*, "Evidence for heritability of abdominal aortic calcific deposits in the Framingham Heart Study," *Circulation*, vol. 106, no. 3, pp. 337–341, 2002.
- [17] J. T. Schousboe, B. C. Taylor, D. P. Kiel, K. E. Ensrud, K. E. Wilson, and E. V McCloskey, "Stroke in Older Women," vol. 23, no. 3, 2008.
- [18] B. K. E. Wilson, "Detection of Abdominal Aortic Calcification with IVA," 2006.
- [19] T. C. Barascuk, N., Ganz, M., Nielsen, M., Register, "Abdominal aortic calcification quantified by the Morphological Atherosclerotic Calcification Distribution (MACD) index is associated with features of the metabolic syndrome.," *BMC Cardiovasc. Disord.*, vol. 11, no. 1, p. 75, 2011.
- [20] M. De Bruijne, "A Pattern Classification Approach to Aorta Calcium Scoring in Radiographs," pp. 170–177, 2005.
- [21] L. A. Conrad-hansen and M. De Bruijne, "Quantizing Calcification in the Lumbar Aorta on 2-D Lateral X-Ray Images," pp. 409–418, 2005.
- [22] K. Elmasri, Y. Hicks, X. Yang, X. Sun, R. Pettit, and W. Evans, "Automatic Detection and Quantification of Abdominal Aortic Calcification in Dual Energy X-ray Absorptiometry," *Procedia Comput. Sci.*, vol. 96, pp. 1011–1021, 2016.
- [23] A. Mohammad *et al.*, "Vertebral fracture assessment-detected abdominal aortic calcification and cardiovascular disease in rheumatoid arthritis.," *Semin. Arthritis Rheum.*

- vol. 43, no. 5, pp. 632–7, Apr. 2014.
- [24] M. H. Horng, “Multilevel minimum cross entropy threshold selection based on the honey bee mating optimization,” *Expert Syst. Appl.*, vol. 37, no. 6, pp. 4580–4592, 2010.
- [25] C. E. Honeycutt and R. Plotnick, “Image analysis techniques and gray-level co-occurrence matrices (GLCM) for calculating bioturbation indices and characterizing biogenic sedimentary structures,” *Comput. Geosci.*, vol. 34, no. 11, pp. 1461–1472, 2008.
- [26] I. Guyon, “An Introduction to Variable and Feature Selection 1 Introduction,” vol. 3, pp. 1157–1182, 2003.
- [27] O. D. Mining, “Orange Data Mining Library Documentation,” 2017.
- [28] J. Demšar *et al.*, “Orange: Data Mining Toolbox in Python,” *J. Mach. Learn. Res.*, vol. 14, p. 23492353, 2013.
- [29] S. Yamada, K. Hashimoto, H. Ogata, Y. Watanabe, M. Oshima, and H. Miyake, “Calcification at orifices of aortic arch branches is a reliable and significant marker of stenosis at carotid bifurcation and intracranial arteries,” *Eur. J. Radiol.*, vol. 83, no. 2, pp. 384–90, Feb. 2014.
- [30] M. Davies and J. L. Fleiss, “Measuring Agreement for Multinomial Data Author (s): Mark Davies and Joseph L . Fleiss Published by : International Biometric Society Stable URL : <http://www.jstor.org/stable/2529886> REFERENCES Linked references are available on JSTOR for this article :,” vol. 38, no. 4, pp. 1047–1051, 2016.
- [31] M. Nelitz, K. P. Guenther, S. Gunkel, and W. Puhl, “Reliability of radiological measurements in the assessment of hip dysplasia in adults,” *Br. J. Radiol.*, vol. 72, no. APR., pp. 331–334, 1999.
- [32] V. M. Patro and M. R. Patra, “Augmenting Weighted Average with Confusion Matrix to Enhance Classification Accuracy,” no. August 2014, 2014.
- [33] E. Honkanen *et al.*, “Abdominal aortic calcification in dialysis patients: results of the CORD study,” *Nephrol. Dial. Transplant*, vol. 23, no. 12, pp. 4009–15, Dec. 2008.
- [34] J. T. Chow, S. Khosla, L. J. Melton, E. J. Atkinson, J. J. Camp, and A. E. Kearns, “Abdominal aortic calcification, BMD, and bone microstructure: a population-based study,” *J. Bone Miner. Res.*, vol. 23, no. 10, pp. 1601–12, Oct. 2008.

Virtual Mouse Control through Natural User Interface

Ayoub Muftah¹, Amna Elhawil², Youssef Omran Gdura³

^{1,2,3}Department of computer Engineering, Collage of Engineering, University of Tripoli,
Tripoli, Libya

¹ayoubmk111@gmail.com, ²a.elhawil@uot.edu.ly, ³y.gdura@uot.edu.ly

ABSTRACT

Keywords:

Human computer interaction, Hand gesture, Skin detection, Image processing

Natural interaction systems have become widely adopted by various applications such as in human computer interaction and virtual environment. In this paper, we present a vision-based gesture application using finger gesture recognition to interact with devices such as computers and tablets without any physical contact. This application was designed to be used as alternative approach to standard mice and keyboards by accepting inputs in a form of hand gesture. The application was implemented using computer vision libraries and existing imaging processing algorithms. The application first detects or recognizes the user's hand, and then performs the required image processing to classify or to distinguish the user's fingers and finally translate all this information to mimic the computer mouse operations. Several real-time tests had been conducted on this application, and the results were satisfactory in terms of design and performance.

I. Introduction

Hand gestures have been in use for a long time in everyday life as a sign language to interact with hear-impaired people [1-3]. Gestures are also used nowadays in Human Computer Interaction (HCI) for communication between people and machines [3-8]. The main goal of HCI is to allow users to control or interact with devices naturally without any mechanical devices such as mice, keyboards, or user interfaces [7-10].

The existing natural interaction technologies can be classified into three basic categories: voice recognition, gesture recognition, or vision recognition (eye-tracking). These technologies are used in different aspects of life such as sign languages translation, voice access to smartphones, voice command while driving, flipping electronic books, and gazing (or eye-tracking) [2] and [11-13]. However, significant amount of research had been conducted on gesture recognition; in particular on hand gestures recognition systems. There are two cases of hand gestures: static or dynamic. Static hand gestures recognition requires the process of a fixed-single image and can be achieved by applying standard pattern recognition techniques [6-9]. In contrast, dynamic hand gesture recognition relies on real-time hand motions, and it is considered a very challenging task in

computer vision [4]. Real-time; dynamic, hand gesture recognition systems can be used for applications like sign language translation, computer interaction, industrial robots, and Virtual Reality (VR) such as games and music [1-3][14] and [15].

Our proposed application implemented a vision-based technology of dynamic hand gesture recognition. The paper is organized as follows: Sections 2 talks briefly about similar related works, whereas Section 3 explains in details the research methodology. Section 3, discusses and analyzes the achieved results. Finally, Section 4 presents the conclusion.

II. Related Work

The first hand gesture recognition system was introduced in 1977 by the Electronic Visualization Lab using a data glove and a number of flex sensors [12]. Since then the field of HCI has seen continuous development of vision-based hand gestures as a new form of natural interaction between human and computer devices [1-8].

In the last two decades, researchers' attention has focused on vision-based dynamic hand gesture recognition techniques because they can potentially provide the most natural and convenient method for interaction in HCI [13 -15]. The vision-

based recognition process can be divided into four main steps: hand gesture detection, then hand tracking and feature

extraction, and finally classification to reach the output gesture [4][5].

Vision-based techniques were used as a computer music controller [2][11]. The authors in [2][11] presented similar solutions to control music player using hand gestures via a computer's webcam. They developed detection system to detect hands and recognize gestures that could trigger different actions such as Windows media player. The vision-based approaches however raise some possible difficulties in regard to the environmental surroundings such as backgrounds and illumination variations which could affect the accuracy of recognition as addressed in [6] and [15-17].

Modern Real-time hand gesture recognition systems tend to highly rely on Artificial Neural Network (ANN) which gave better results. Although the neural network requires a large dataset to train the system, the Convolutional Neural Networks (CNN) and Complex-Valued Neural Network (CVNN) architectures provide very good average recognition accuracies either for static or real-time hand gesture detection [7 - 10]. The purpose of these techniques is to implement sophisticated systems for human-computer communication. Computer mouse control via hand gesture is an example. In this field, many studies have considered this important application using different techniques. In [18] an external webcam captures the image of the hand put on a pad. The fixed image is processed and color detection are used for gesture interpretation. In [19] the gesture recognition is performed using the background subtraction method. A valuable literature survey on 10 vision-based hand gesture systems and their techniques and methodologies is given in [13].

III. Research Methodology

The analysis phase is considered to be one of the most important stages especially in the technologies that may enter in a constructive matter. In this work, the analysis consists of several basic stages:

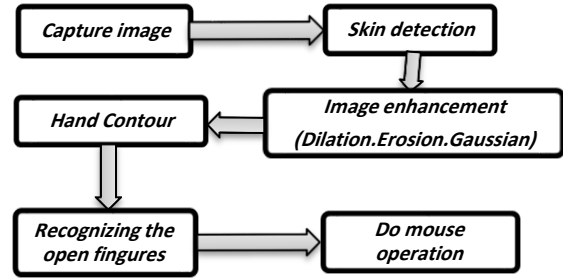
- a) Capturing the frame
- b) Skin detection
- c) Image enhancement
- d) Extracting hand contour
- e) Recognizing the open fingers
- f) Perform the analyzed operation

All these stages are summarized in the block diagram depicted in Fig. 1. In the next sections we explain in details each stage.

Fig. 1. Block diagram of analysis steps

A. Capturing the frame

The capture frame step is responsible for fetching camera frames in real time such that each frame is processed



separately.

B. Skin detection

In this stage the hand gesture in a frame is detected. The most common techniques for skin detection are Hue Saturation Value (HSV) and YCbCr [16] and [20 – 22]. In the studies [21 - 22] both techniques retain mostly same good performance. However, the skin detection using HSV was successful in [20] while in [16] HSV results were not satisfactory. In this work we implemented both techniques and evaluated them in order to choose the most appropriate one for our application.

a) HSV (Hue,Saturation, Value)

Hue is the effective color of a given area and *Saturation* represents the bright whereas the “*value*” is the color luminance [19]. The HSV algorithm can be summarized in the following steps [16]:

- 1) Convert the RGB image to HSV components. The simplest way to do that is by changing the range of R, G and B values from [0, 255] to [0, 1] as following [16]:

$$\hat{R} = R/255 \quad (1)$$

$$\hat{G} = G/255 \quad (2)$$

$$\hat{B} = B/255 \quad (3)$$

$$C_{max} = \max(\hat{R}, \hat{G}, \hat{B}) \quad (4)$$

$$C_{min} = \min(\hat{R}, \hat{G}, \hat{B}) \quad (5)$$

$$\Delta = C_{max} - C_{min} \quad (6)$$

Hue calculation:

$$H = \begin{cases} 60^\circ * \left(\frac{G'-B'}{\Delta} \% 6 \right) & \text{if } C_{max} = \hat{R} \\ 60^\circ * \left(\frac{B'-R'}{\Delta} + 2 \right) & \text{if } C_{max} = \hat{G} \\ 60^\circ * \left(\frac{R'-G'}{\Delta} + 4 \right) & \text{if } C_{max} = \hat{B} \end{cases} \quad (7)$$

Saturation calculation:

$$S = \begin{cases} 0 & , C_{max} = 0 \\ \frac{\Delta}{C_{max}} & , C_{max} \neq 0 \end{cases} \quad (8)$$

Value calculation:

$$V = C_{max} \quad (9)$$

- 2) The pixel is considered to be a skin pixel if the corresponding saturation values are in the range of (0.20, 0.75), the value exceeds 0.35 and its hue lies between (0, 25) [16].

Fig. 2. shows sample of the results of the HSV technique. Unfortunately, the technique fails to detect the hand.

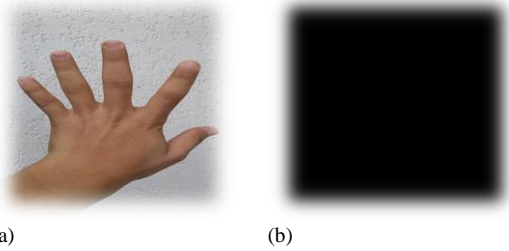


Fig. 2. a) Original image. b) The result of HSV Skin detection algorithm

b) YCbCr (Luma, Blue Chroma, Red Chroma) [19]

YCbCr is an encoded nonlinear RGB signal. It is commonly used in video and digital photography systems. Y is the luma component, C_b is the blue component and C_r is the red-difference chroma. The conversion from RGB to YCbCr color space can be accomplished by following matrix:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (10)$$

For the skin pixel using YCbCr: red chrominance value lies between 140 and 165, blue chrominance value lies between 140 and 195 and hue value is in the range of 0.01 to 0.1 [16]. A sample of obtained result is shown in Fig. 3. We have applied many tests and the success of this technique was very obvious. However, we have considered it in this work.

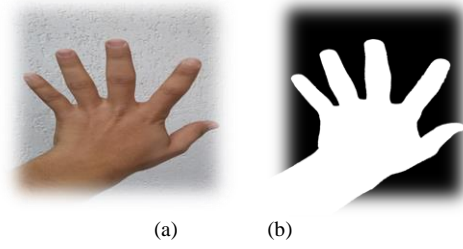


Fig. 3. a) Original image. b) The result of YCrCb algorithm

C. Image Enhancement

After skin detection, it is required to filter the extracted image and remove the noise. This process is accomplished by eroding and dilating the image. That helps to make the boundaries more shapers and clear. The noise are reduced via smooth Gaussian blur effect. In this paper, we used 3 x 3 masks for erosion process and 5 x 5 masks for dilation. All unnecessary details in the hand image have been removed.

D. Extracting Hand Contour

The extraction of the hand contour or the region of interest (ROI) is performed using OpenCV's *findContours* function. This function is based on Moore neighborhood algorithm. Each pixel, also called pattern, is considered as a center of two-dimension square matrix with 8 neighbors surrounding it as shown in Fig. 4. The idea is to scan these neighboring pixels starting from the leftmost column to the right and from the bottom to the top in the clockwise direction. All detected black pixels are considered as the contour of the pattern.

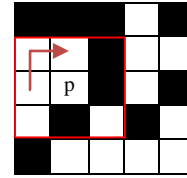


Fig. 4. Moore Neighborhood algorithm, p is the pattern

E. Recognizing the Open Fingers

The number of open figures needs to be recognized from the hand image extracted from the previous steps. For this purpose, we applied convex hull (CH) and convex defect operations. The first operation is used to extract defects which will be intersect between hand contour and hand convex hull. CH function in OpenCv is used to close the gabs (holes) in the shape and convert it to a closed shape. On the other hand, convexity defects of the points that represent the outline of a hand give clues as to the location of the finger tips.

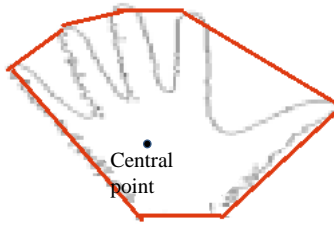


Fig. 5 Hand convex hull and the hand center

The next step is to determine the central point. This can be done by either determining the smallest rectangle area of the shape (hand), as shown in Fig 5, then calculating the central point of the rectangular [23] or alternatively by using a center tracking unit as suggested by [24]. In this work, the first approach is implemented.

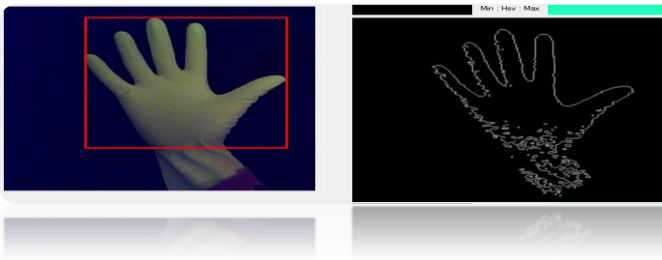


Fig. 6 Getting the minimum area

The last step is to extract all hand defects, as previously mentioned.

- The defect is instructed between CH and contour as shown in Fig. 6.
- The defect is compound from start point, depth point and end point as depicted in Fig. 7.
- The end point will be the start point of the adjacent finger.

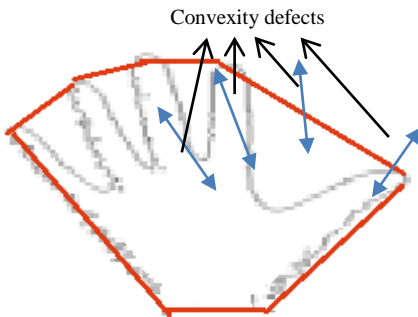


Fig. 6 Hand defects

After the extraction step, defects are filtered using some conditions. This process helps to determine how many fingers are open in the hand. These conditions are:

1. The y coordinate of the depth point must be greater than that of the start point ($y_d > y_s$).
2. The center of hand must be greater than both the depth and the start point.
3. The distance between the depth and the start point ($x_d - x_s$) must be greater than the relative height of shortest finger in hand.

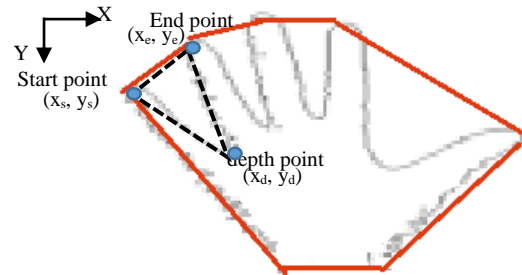


Fig. 7 Determining the open fingers

F. Control Mouse Based On Analysis Frame

Based on the resulted image obtained from the previous steps, the mouse control is performed. The control includes either moving the mouse pointer to the coordination of the detected hand or making an active click based on the number of open fingers.

IV. Results

This work is implemented using OpenCV C++ in Visual Studio 2017. The basic user interface is depicted in Fig. 8.

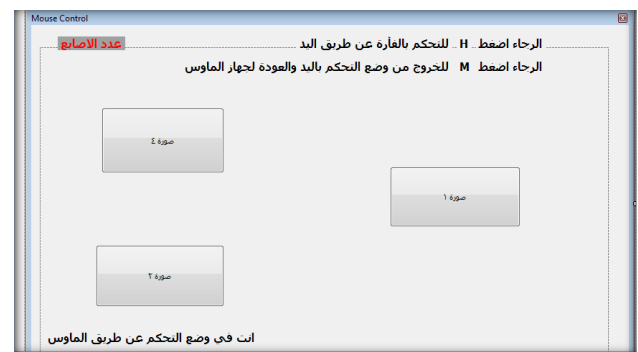


Fig. 8 Basic user interface

The detected number of open fingers is displayed in a label placed on the top-left of the window. Figs. 9 and 10 show the RGB and gray scale images respectively. In Fig. 10, it is obvious the success of the skin recognition for both the hand and the face of the user. In addition, Fig. 10 shows the obtained results when the user opens his 4 fingers. In addition,

Fig. 11 illustrates the obtained results of thumb finger detection. Once the user closes his hand the detected open fingers is zero as shown in Fig. 12. In this work, the mouse pointer moves based on the hand movements. The cursor follows the central of the hand. On the other hand, the left and right mouse click events are done based on the one and two open fingers respectively.

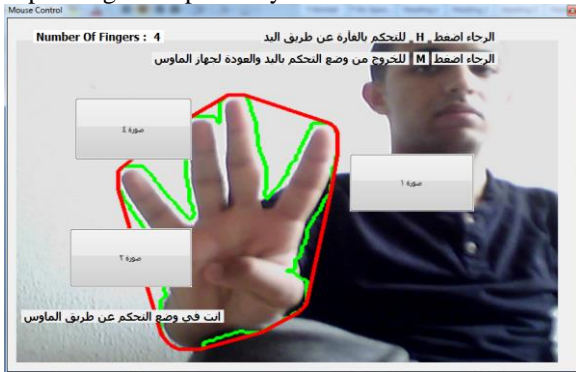


Fig. 9 Captured RGB resulted image



Fig. 10 Captured black and white image of the image in Fig. 9

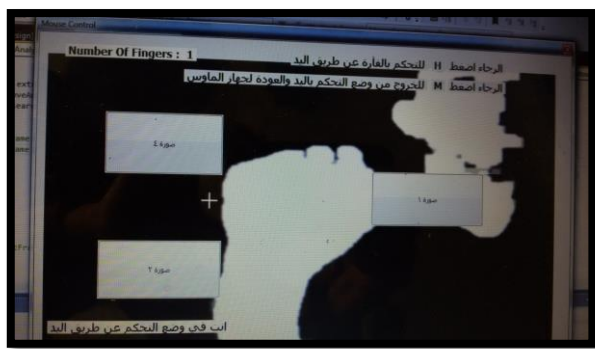


Fig. 11 One open finger detection

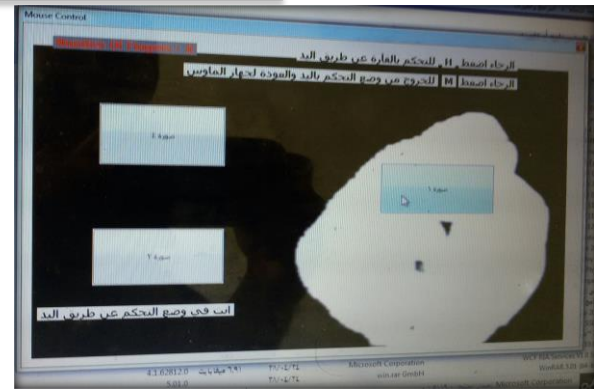


Fig. 12 Zero open finger detection

V. Conclusion

The system is built to recognize gestures and trigger the specified events that correspond to computer-mouse functions. The results are very motivating and can be exploited to control a robot or to interact with interface parts of the computer such mouse or keyboard.

References

- [1] K. Hoshino, "Hand Gesture Interface for Entertainment Games", Handbook of Digital Games and Entertainment Technologies, 2016, pp. 293-312
- [2] J. Jiang, J. Ma and Y. Jin, "Computer Music Controller Based on Hand Gestures Recognition Through Web-cam", EE368 - Digital Image Processing, Stanford University, 2012.
- [3] R. Zaman and N. Abraheem, "Comparative Study of Hand Gesture Recognition System", Computer Science & Information Technology, 2012, vol. 2. pp203-213.
- [4] S.. K. Yewale and P. K. Bharné, "Hand gesture recognition using different algorithms based on artificial neural network", 2011 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), April 2011.
- [5] P. Parvathy, K. Subramaniam, P. Venkatesan, P. Karthikaikumar, J. Varghese and T. Jayasankar, "Development of hand gesture recognition system using machine learning", Journal of Ambient Intelligence and Humanized Computing (2020), <https://doi.org/10.1007/s12652-020-02314-2>
- [6] A. B. Jmaa , W. Mahdi, Y. B. Jemaa and A. B. Hamado, "A New Approach For Hand Gestures Recognition Based on Depth Map Captured by RGB-D Camera", Comp. y Sist. [online]. 2016, vol. 20, no.4, pp.709-721.
- [7] O. Köpüklü, A. Gunduz, N. Kose and G. Rigoll, "Real-time Hand Gesture Detection and Classification Using Convolutional Neural Networks", IEEE International Conference on Automatic Face and Gesture Recognition (FG 2019), Oct 2019
- [8] G.R.S. Murthy and R.S. Jadon, "Hand Gesture Recognition using Neural Networks", 2010 IEEE 2nd International Advance Computing Conference (IACC), Feb. 2010.
- [9] A. R. Hafiz, M. F. Amin, and K. Murase, "Real-Time Hand Gesture Recognition Using Complex-Valued Neural Network (CVNN)", International Conference on Neural Information Processing ICONIP 2011, 2011, pp 541-549.
- [10] H. Lin, M. Hsu, W. Chen, "Human hand gesture recognition using a convolution neural network", 2014 IEEE International Conference on Automation Science and Engineering (CASE), 2014.
- [11] WAS NO 3 A. R. Parab, J. R. Ghosalkar, V. Akshay, B. Medge and M. Meena, "Music Controller using Gesture Recognition to Control Music Playback", National Conference on Technological Advancement & Automatization in Engineering, Jan. 2016.

- [12] D. J. Sturman and D. Zeltzer, "A survey of glove-based input," in IEEE Computer Graphics and Applications, vol. 14, no. 1, pp. 30-39, Jan. 1994.
- [13] S. Pranit, K. Pandya, S. Harsh, G. Jay, "Survey on Vision based Hand Gesture Recognition. International Journal of Computer Sciences and Engineering, 2019, Vol. 7. 281-288.
- [14] Y. Zhu, Z. Yang and B. Yuan, "Vision Based Hand Gesture Recognition," 2013 International Conference on Service Sciences (ICSS), Shenzhen, China, 2013, pp. 260-265.
- [15] M. Yassen and S. Jusoh, "systematic review on hand gesture recognition techniques, challenges and applications", PeerJ Computer Science 5:e218 <https://doi.org/10.7717/peerj-cs.218> 2019.
- [16] G. P. Surampalli, D. J and D. M "An Analysis of Skin Pixel Detection using Different Skin Color Extraction Techniques", International Journal of Computer Applications (IJCA), vol. 54, No. 17, pp. 1-5.
- [17] M. Deswal and N. Sharma, "A Fast HSV Image Color and Texture Detection and Image Conversion Algorithm", International Journal of Science and Research (IJSR), vol. 3 Issue 6, June 2014.
- [18] H. Grif and C. C. Farcas, " Mouse Cursor System Based on Hand Gesture", Procedia Manufacturing 22:1038-1042, Jan. 2018.
- [19] Z. Chen, J. Kim, J. Liang, J. Zhang and Y. Yuan, "Real-Time Hand Gesture Recognition Using Finger Segmentation", The Scientific World Journal, vol. 2014.
- [20] V. A. Oliveira and A. Conci, "Skin Detection using HSV color space", International Journal of Computer Trends and Technology, 2009.
- [21] K. Nikolskaia, N. Ezhova, A. Sinkov, and Maksim Medvedev, "Skin Detection Technique Based on HSV Color Model and SLIC Segmentation Method", Proceedings of the 4th Ural Workshop on Parallel, Distributed, and Cloud Computing for Young Scientists Yekaterinburg, Russia, November 15th, 2018.
- [22] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat and J. Jatakia, " Human Skin Detection Using RGB, HSV and YCbCr Color Models", Advances in Intelligent Systems Research, vol. 137, pp. 324-332, 2017.
- [23] R. Hartanto, A. Susanto and P. I. Santosa, "Real Time Hand Gesture Movements Tracking and Recognizing System", The 7th – Electrical Power, Electronics, Communications, Controls, and Informatics International Seminar 2014, Indonesia., 2014.
- [24] A. V. W. Smith, A. I. Sutherland, A. Lemoine and S. Mcgrath, " Hand gesture recognition system and method", Patent US 6128003 A, 2000.

R-ORT: A Robust Object Recognition Technique from 2D Images Regardless the Viewing Angles

Rawia O. Almontaser ¹, Omar Mabrok Bouzid ², Ayad Ali Keshlaf ³

^{1,2} Electrical and Electronic Engineering Department Faculty of Engineering, University of Gharyan, Gharyan, Libya

³ Computer Department Faculty of Engineering, Sabratha University
Sabratha, Libya

¹ rawia9620@gmail.com, ² omar.bouzid@gu.edu.ly, ³ ayad.Keshlaf@sabu.edu.ly

ABSTRACT

Keywords:

3D object recognition, corner detector, feature extraction, recognition accuracy, bounding box area.

Accuracy is a very important issue for a wide range of objects recognition and identification applications, including robotics, industrial automation systems, and content-based image retrieval processes. In this context, this paper presents a 3D object recognition system called R-ORT that recognizes and identifies objects (namely: a pyramid, a cubic, and a sphere) based on their 2D images regardless of the fact that objects seen in novel viewing angles negatively affect the recognition accuracy. The system imports an object image via a webcam, and extracts an object contour from a Canny edge map and then it recognizes the object type based on the number of detected contour corners. Although every object has a certain number of corners and differs from the other, the preliminary evaluation results show that in some cases the pyramids and cubes can have the same number of corners due to variations in viewing angles of the objects. To make the object recognition independent from such effect, the system uses besides the corner features the ratio of the object's area to the area of its bounding box as combined features. Experimental results of this study illustrated that the recognition accuracy is raised from 83% (in the case of using only the corner features) to 100% when the combined features are utilized under the condition of good lighting.

I. Introduction

One of the hot research areas in the field of computer vision and image processing is objects recognition and identification. This is because of their applications are found in various aspects of computer vision tasks [1]. Object recognition can be defined as a process of recognizing each object in a given image, while the identification process is the procedure of labeling these objects with their associated names based on predefined features [2]. In fact, it is easy for a human being to accurately do the object recognition and identification for real-world objects, however, it would be a difficult task, if a robotic or an automated system is intended to be used, unless a certain percentage of errors is acceptable.

In this aspect, the variations in viewing angles of the objects is a real challenge, which should be considered and tackled to increase the accuracy rate of the object recognition and identification processes. These processes are based on the choice of appropriate features that are the most closely to the detected object. In fact, there are many types of objects (i.e. geometry shapes, table, chair, car, bicycle, a person, birds, or any real thing) and each of them can be recognized either by one feature or by a combination of various features. The features could be color [3], texture [4], shape [5], edges [6], and corners [7]. Yet, to increase the recognition and identification accuracy, the selection of appropriate features should depend on the object type.

The rest of the paper is organized as follows. Section II reviews the related work to object recognition, while Section III introduces the main steps of R-ORT. Tackling the viewing angle challenge will be the topic of Section IV. Section V will give the details of the R-ORT prototype. In Section VI, the results of the proposed system are presented and discussed. Finally, conclusions and future work are given in Section VII.

II. Related Work

Objects recognition and identification is the greatest importance for various automation and robotics applications. Literary, this is a well-known topic in computer vision and there are many researchers worked on this topic. The study in [2] reviewed the state-of-the-art work of object recognition techniques, which covered numerous feature extraction methods and classification techniques that are used for this purpose. The study concluded that although there are tremendous of object recognition works but, the accuracy of many recognition results are still low.

In [8], various objects were recognized and sorted in real-time based on Bag features. However, recognition accuracy was affected by lighting conditions and unclear backgrounds. In [9], an automatic sorting system based on a computer vision was proposed, this system was tended to use for sorting parcel boxes via calculating their volumes with an accuracy of 87.5%. In [10], the author proposed an object recognition system based on the use of K-Nearest neighbor method. This method is supported by the utilization of Eigen values extracted from the features of an image. Authors of reference [11] proposed a new algorithm to detect and recognize 2D shapes from an image.

They claim that this algorithm has the ability to recognize most of the shapes.

In [12], features such as color, shape and texture have been used to classify objects using various image processing methods. The proposed automatic sorting system received frames of images to process them and detect the specified features. However, the rotation of the object- due to the change in its angles of inclination- affects the 2D shape recognition (i.e. decreases the recognition accuracy). This leads to utilize the object area in addition to the area of its bounding box in order to find the ratio between both areas which helped in tackling such challenge.

Reference [7] presents a system that recognized and identified three types of 3D objects using corner features. The experimental results of this system showed that recognition errors were caused due to the light effect and the angle view of the camera, and they reached 6% as a maximum. The authors of this study did not take into their account the “viewpoint-dependent theories” [13], which mean the object viewpoint is changeable based on the viewpoint at which it is seen and consequently the recognition process will be prone to errors and imprecisions. One of the suggested solutions is to store each viewpoint and angle of the individual objects, but this form of recognition requires a big size of memory.

In the context of this literature review, we found that the closest related study to our work is the work presented in [7]. This work does not provide a solution for making the object recognition independent of its rotation. Therefore, the main aim of the proposed R-ORT is to introduce a simple, not costly, and easy to implement solution that can tackle such a challenge so that the recognition accuracy for sorting 3D objects using image processing will be improved. This will be achieved via the integration of both corner features and the area ratio between the area of the object and the area of its bounding box, as seen in the next sections.

III. R-ORT Steps

The **R-ORT** consists of five fundamental steps, namely: image acquisition, image enhancement (pre-processing), object segmentation, features extraction, object recognition. These steps are described as follows.

B. Image Acquisition

In the first stage, images of objects are captured in a true color format in real-time from an incoming video stream. In the proposed system, the camera itself is assumed to be stationary and objects are randomly located in front of the camera. This means the object viewpoints are changeable with respect to the camera view.

C. Image Enhancement

In order for object images to become ready for use as well as for basic processing, they must be first pre-processed. This is composed of two basic operations.

B.1 Converting Color Images into Grayscale

There are many types of images encoded by digital computers, which are saved as arrays of pixels. The most commonly used types of images are color, grayscale, and binary. Each component (pixel) in the color image consists of three values indicating the intensity of the color at that point, while each element in the grayscale image consists of only one value indicating the intensity of the gray color [14]. This is the reason behind using grayscale images in the **R-ORT**.

B.2 Image Filtering

Image filtering is considered as one of the basic and primary steps for image processing to filter noise and other defects in digital images [15]. Within the **R-ORT**, the Gaussian method was used to filter the images of objects with specific characteristics summarized in Table I.

TABLE II. THE CHARACTERISTIC OF USED GAUSSIAN FILTER.

Filter type	Low pass filter
Standard deviation of the Gaussian distribution σ	0.84089642
Filter size	7×7
Filter domain	Spatial domain

D. Object Segmentation

After the filtering process, the process of segmenting the image pixels into meaningful structures is launched, which known as a segmentation process [16]. This process specifies the region of interest in the image (where an object appears). More precisely, to isolate pixels belonging to this object (i.e. most important information) from the rest of the image, so that the required processing can be carried out on these pixels to extract relevant features.

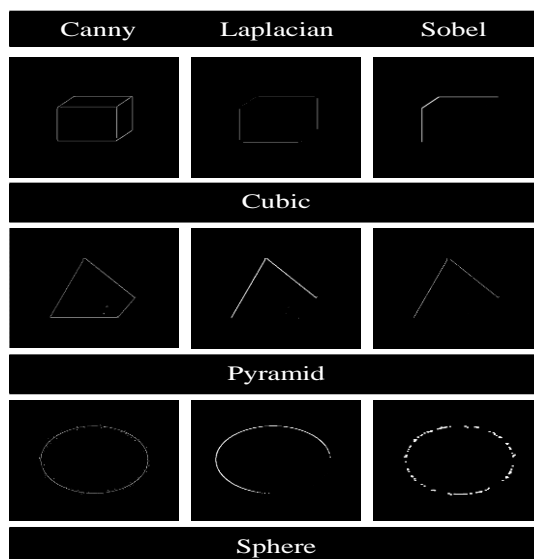


Figure 1 Result of various edge detection approaches applied on three types of objectives.

Three types of algorithms are used for edge detection: Canny, Laplacian, and Sobel. These algorithms have a variation in their results when applied to extract edges from the same image. Therefore, their effect was studied in order to specify the best algorithm among them. Figure 1 illustrates the effects of using these algorithms, through which it can be seen that the use of the Canny method is the best in extracting all edges of the object, thus it is adopted by the **R-ORT**. This conclusion is also consistent with the findings of the work presented in [17].

E. Feature Extraction

The consequence of applying an edge detection method to an image is to filter out irrelevant information (i.e. reducing the amount of data to be handled), at the same time it preserves important fundamental features of an image. There are different classes of feature extraction methods that can be used, such as feature-based and shape-based methods [2], both of them are used by the **R-ORT**.

F. Object Recognition.

Object recognition is the final step of the **R-ORT**, which leads to distinguish and separate objects based on the features extracted in the previous step. The predicted output of this stage is fed to the hardware part of the proposed automated system in order to sort these objects as discussed next.

IV. Tackling the Viewing Angle Challenge

A real problem related to viewing angle challenge was faced during the evaluation of the **R-ORT**. This is clear when all object corners are calculated. It has been found from the preliminary results that the pyramid, the cubic, and the sphere have in most cases 3, 6, and 8 corners, respectively. However, in some cases, the pyramid and the cubic could have 4 or 5 corners due to "viewpoint-dependent theories", as shown in Figure 2. As seen in this figure, although Figures 2a and 2b represent different objects, both figures have the same number of corners (5). The same thing is valid for Figures 2c and 2d where the number of corners is 4. This makes distinguishing between both of them very difficult.

In order to solve this problem and tackle the viewing angle challenge, the ratio of an object area to the area of its bounding box is calculated, as illustrated in Equation (1).

$$\text{Area ratio} = \text{Object area} / \text{Bounding box area} \quad (1)$$

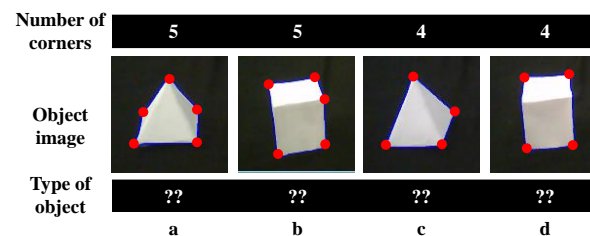


Figure 1. Effect of "viewpoint-dependent theories" on object recognition.

Experimentally and as shown in Figure 3, it has been found that if the ratio between the two areas is less than 80%, then the object is classified as a pyramid, but if this percentage is greater than that, the object is classified as a cubic. Figure 3 clarifies this calculation, which is helped in tackling the previous challenge. In this regard, it is important to mention that both areas are calculated in pixels. The flowchart of the *R-ORT* is presented in Figure 4.

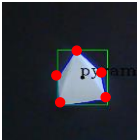
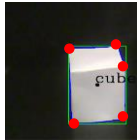
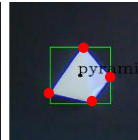
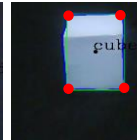
# of corners 5	# of corners 5	# of corners 4	# of corners 4
			
Object area (pixel)			
30403	49549	33340	50340
Bounding box area (pixel)			
45881	52287	47250	50264
Ratio of two areas			
66%	95%	70%	99%

Figure 2. Finding the percentage threshold of the area ratio.

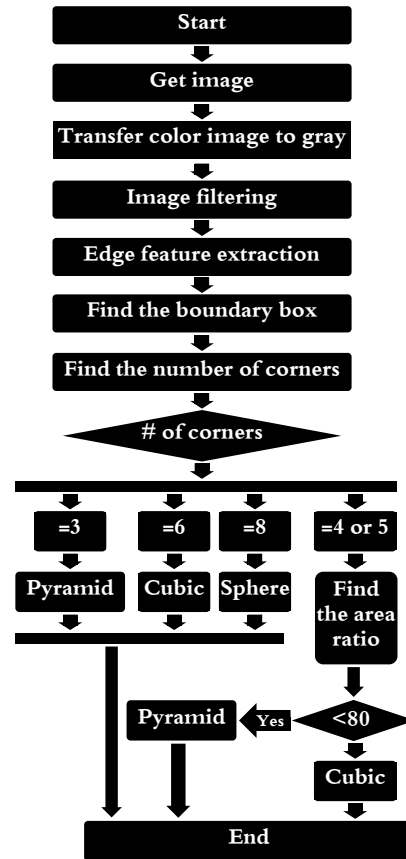


Figure 3. The main flowchart steps of the *R-ORT*.

V. R-ORT Prototype

This section introduced a prototype that has been built based on *R-ORT*. This prototype is designed and implemented as a proof-of-concept model for this technique. As shown in Figure 5, the prototype is divided into three stages. The first stage represents the input part which senses the physical objects. In the second stage, the acquired images are processed and relevant features are extracted as discussed in the previous sections. Finally, the results are fed to the last stage by means of the dedicated software to control a conveyor belt using a microcontroller and a DC stepper motor.

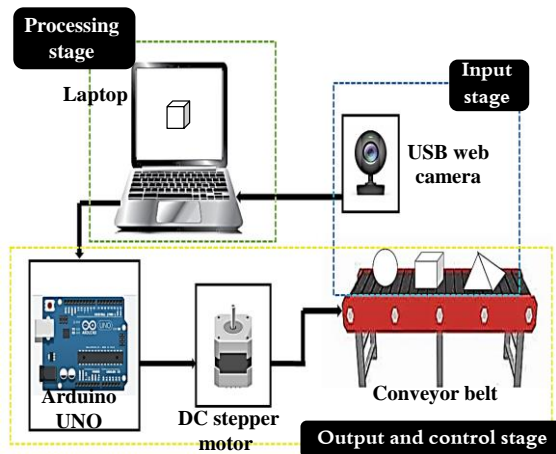


Figure 4. The proposed automated real-time 3D object recognition system.

The hardware prototype of the proposed system comprises hardware components such as an ordinary laptop, a simple USB web camera, an Arduino UNO board, a DC stepper motor and a conveyor belt. The experiments were conducted in an ordinary indoor environment. And several experiments were conducted to classify test objects.

VI. Results and Discussion

The 3D object recognition system presented in this study was tested by running several experiments on the prototype designed for the purpose of knowing and evaluating the performance of this system. Three objects with different sizes (shown in Figure 6) including a pyramid, a cube, and a sphere were used in these experiments.

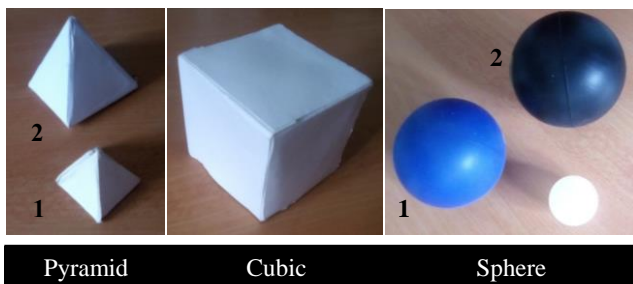


Figure 5. Three types of test objects with different sizes.

The experiments were conducted under good conditions of illumination. The distance between the object and the camera was between 30 to 40cm, the position of the camera was fixed and objects were randomly positioned, so that objects were seen in novel viewpoints at each trial. Figure 7 depicts the results of the experiments, which reflect the performance of the *R-ORT*. As can be seen in this figure, all objects have been detected and recognized despite the fact that they have different viewpoints. This was due to the integration of both corner features and the area ratio between the object area and the bounding box area, which increases the recognition accuracy from 85% to 100%.

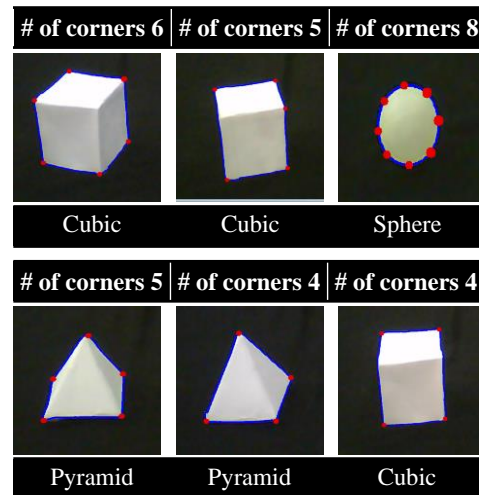


Figure 6. Examples of recognition of test objects with 100% accuracy.

Table 1 summarizes the recognition accuracy of the conducted experiments for the proposed technique. These results were for the both cases: using both features and using only the corner feature. As Table 1 illustrated, this technique was able to identify all the objects with 100% accuracy for the former case, while for the later case the recognition accuracy has an average of 85%.

TABLE II
RECOGNITION ACCURACY OF TEST OBJECTS WITH DIFFERENT VIEWING ANGLES

Object type	Number of experiments	Accuracy using corner feature	Accuracy using both features
Sphere 1	9	100%	100%
Sphere 2	11	100%	100%
Cubic	20	71%	100%
Pyramid 1	8	79%	100%
Pyramid 2	9	74%	100%
Average	-	85%	100%

In the context of these experiments, the effect of the image resolution on the recognition accuracy was tested. It was found that changing the camera resolution settings to the value of 82x83, 125x126, and 248x251 does not affect this accuracy since all objects were detected and recognized under these settings.

Experiments were also carried out under different lighting conditions (namely good and bad). The results of these experiments showed that as the lighting condition becomes low, the recognition accuracy becomes worse as shown in Figure 8. This is because of the closeness of the chromatic values of the image resulted under the bad lighting condition. This convergence resulted in no significant difference between the

values representing edges and the values representing features in the image.

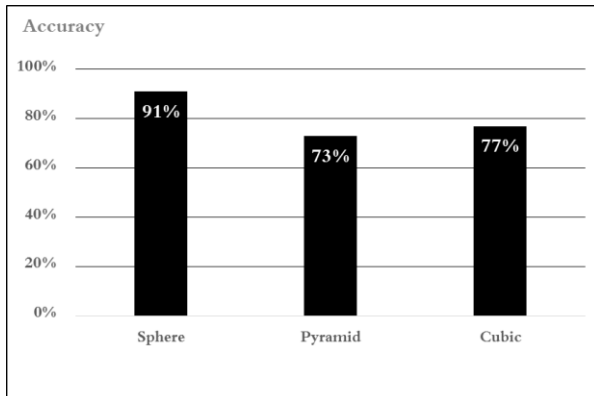


Figure 7. Recognition accuracy under bad lighting conditions.

VII. Conclusion and Future work

Object recognition is viewed as a crucial nevertheless complicated task that must be efficiently executed and hopefully with high accuracy. Thus, researchers around the world are still trying to improve the accuracy of the recognition process. This paper introduced a technique called **R-ORT** for objects recognition and identification. This technique uses 2D images to recognize three types of objects, namely: a pyramid, a cubic, and a sphere with a high accuracy. The high accuracy was gained through the integration of the corner features with the ratio between object area and its bounding box area. Both features helped in tackling the effect of “viewpoint-dependent theories”, particularly for both pyramid and cubic objects. Experimentally, we found that if the ratio between the two areas is less than 80%, then the object is classified as a pyramid, but if this percentage is greater than that, the object is classified as a cubic. Generally, experiments results illustrated that the **R-ORT** has identified all the testing objects with an accuracy of 100% in a good lighting condition. The future work of this study will focus more on the effect of varying illumination settings on recognition accuracy.

References

- [1] Bansal, M., Kumar, M., Kumar, M. et al. “An Efficient Technique for Object Recognition Using Shi-Tomasi Corner Detection Algorithm,” *Soft Computer*, 2020.
- [2] Bansal, M., Kumar, M. & Kumar, M., “2D Object Recognition Techniques: State-of-the-Art Work,” *Archives of Computational Methods in Engineering*, 2020.
- [3] Qiu, G., Sudirman, S., “A Binary Color Vision Framework for Content-Based Image Indexing,” in: *proceedings of the international conference on advances in visual information systems*, pp. 50–60, 2002.
- [4] Wu, X., Bhanu, B., “Gabor Wavelet Representation for 3-D Object Recognition,” *IEEE Trans. Image Process* 6(1), pp. 47–64, 1997.
- [5] Belongie, S., Malik, J., Puzicha, J., “Shape Matching and Object Recognition Using Shape Contexts,” *IEEE Trans. Pattern Anal Mach Intell* 24(4), pp. 509–522, 2002.
- [6] Kumar, I., Rawat, J., Bhaduria, H.S., “A Conventional Study of Edge Detection Technique in Digital Image Processing,” *Int J Computer Sci Mobile Computer* 3(4), pp. 328–334, 2014.
- [7] Kitti, T., Jaruwan, T., Chaiyaporn, T., “An Object Recognition and Identification System Using the Harris Corner Detection Method,” *Int J Mach Learn Computer* 2(4), pp. 462–465, 2012.
- [8] Rahul Vijay Soans, Pradyumna G. R, Yohei Fukumizu, “Object Sorting Using Image Processing,” *3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 2018.
- [9] Riky Tri Yunardi, Winarno, Pujiyanto, “Contour-based Object Detection in Automatic Sorting System for a Parcel Boxes,” *Int’l Conf. on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA 2015)*, 2015.
- [10] Muralidharan, R., “Object Recognition Using K-Nearest Neighbor Supported By Eigen Value Generated From the Features of an Image,” *International Journal of Innovative Research in Computer and Communication Engineering*, 2(8), 2014
- [11] El Abbadi, N., and Al Saadi, L., “Automatic Detection and Recognize Different Shapes in an Image,” *IJCSI International Journal of Computer Science Issues*, 10(6, 1), 2013.
- [12] Sabnis, M., Thakur, V., Thorat, R., Yeole, G., Tank, C., “Object Sorting in Manufacturing Industries Using Image Processing,” *International Journal of Computer Engineering and Applications*, 9(5), 2015.
- [13] Tarr, M., Bulthoff, H., “Is Human Object Recognition Better Described by Geon Structural Descriptions or by Multiple Views? Comment on Biederman and Gerhardstein (1993),” *Journal of Experimental Psychology: Human Perception and Performance*, 21(6), pp. 1494–1505, 1995.
- [14] da Silva, E. A. B. and Mendonca, G. V., “Digital Image Processing”, in *The Electrical Engineering Handbook*, Chapter VII.4, pp. 891-910. Editor: Wai-Kai Chen, Elsevier – Academic Press, 2005.
- [15] Dhar, R., Gupta, R., and Baishnab, K., L., “An Analysis of Canny and Laplacian of Gaussian Image Filters in Regard to Evaluating Retinal Image,” *International Conference on Green Computing Communication and Electrical Engineering (ICGCCCE)*, pp. 1-6, 2014.
- [16] Fisher, R.B., Dawson-Howe, K., Fitzgibbon, A., Robertson, C., and Trucco, E., “Dictionary of Computer Vision and Image Processing,” Wiley, 2005.
- [17] Muthukrishnan, R., Radha, M., “Edge detection techniques for image segmentation,” *Int J Comput Sci Inf Technol* 3(6), pp.259–267, 2011.

Speed Control of DC Motor Using Simulink With Arduino

¹Asma Alfergani, ²Shoroug Alweheshi, ³Abdelhameed Musbah Saeed, ⁴Fathi Adam and ⁵Hussamedin S. Mohamed

^{1,2,3,4,5} Electrical and Electronics Engineering Department, University of Benghazi, Benghazi, Libya

¹ asma.alfergani@uob.edu.ly, ² kadershrik@gmail.com, ³ hameed.elmegricy@gmail.com,

⁴ fbotaher@gmail.com, ⁵ hsbargathy@gmail.com

Keywords:

Arduino; DC motor;
PID; H-Bridge

ABSTRACT

Nowadays, DC machines, in particular motors, are widely employed in most applications. DC machines that convert the energy form from mechanical to electrical are called DC generators, whilst the DC machines that convert the electrical to mechanical are the motors. DC motors currently are the most commonly used due to their high torque at starting and easy maintenance and installation compared to their AC counterparts. Since the speed of the DC motor is crucial to the functioning of the motor, it is important to control it while varying the voltage. Therefore, this paper proposed controlling the speed through controlling input voltage, using a PID controller with a closed loop system applied practically via utilizing Arduino and Matlab/Simulink to generate the control signal. The results of connecting the Feedback system were compared to their counterparts while connecting the open loop system.

I. Introduction

Nowadays, the Direct Current, DC, machines are one of the most that commonly used in the field of industry, precisely the DC motors, due to their low cost, easy maintenance and the simple control techniques applied for speed control [1].

DC motors are becoming a great part of many applications, such as robot arm drives, machine tools, rolling mills, and aircraft control [2].

The DC motor turn electrical energy into mechanical by electromagnetic induction, where the circuit is composed of a stator (the stationary field) and a rotor (the rotating field), the current carried in the armature generates a magnetic flux which through interaction will produce torque and rotational speed [3].

The speed of the motor is controlled through controlling the value of the input voltage, where it can be adjusted according to the load demand. As all the DC motors require a controller to alter the speed to a desired value, the control technique can

be performed manually or by applying some conventional control techniques such as PID control [4]

The crucial concept of the PID controller is tuning the proportional, Integral and Derivative gains, P, I, and D, respectively [4], this technique is widely utilized due to its simple structure and fast response to improve the system's transient and the steady state behavior, as well as obtaining fast settling time without causing overshoot [5,6].

This paper aims at studying the behavior of DC motor coupled with DC generator load, which is derived by H-bridge at required speed. In addition, the PID controller is used to get the best performance practically in the laboratory, where the paper is organized as follows:

The following section is dedicated to investigate the open loop controller system behavior, then a practical implementation of the latter is demonstrated. In the fourth section a closed loop system and its characteristics are illustrated, followed by a comparison the results of both systems. Finally, results of applying PID controller are showcased.

II. Open loop system description

A. Open Loop DC Motor Speed Controller system

In the open loop systems, the signal indicated by the Arduino was sent to the H-Bridge to drive the motor at a certain speed, according to the level of voltage resulted from PWM signal. The two aforementioned DC machines connected to the same shaft, where the first works as a motor converting the electric energy into mechanical whose output is used as an input for the second DC generator and the energy converted is considered as the expressing speed.

For instance, the parameter's pins, the inputs and outputs are defined as IN1, IN2 and pulse width modulation (PWM) signals, respectively. For direction of DC motor, the values of the inputs are assigned as IN1 low and IN2 high, then the "serial monitor" function is called as illustrated in Fig. 1, consequently, a value of the speed is manually appointed from 0 to 255 as it is considered the number of samples compared to PWM signals.

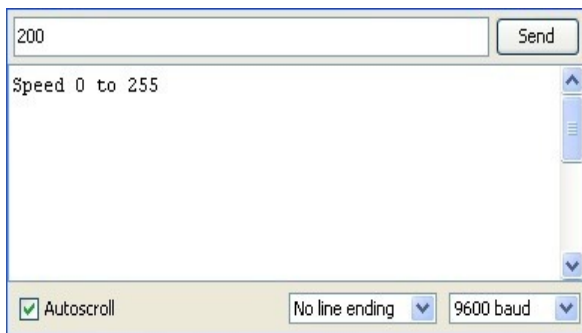


Fig. 1. The serial monitor

G. The Hardware Connections

As demonstrated in Fig. 2, the system applied in this paper comprises of two DC machines connected to the same shaft, where the former works as a motor and the latter as a

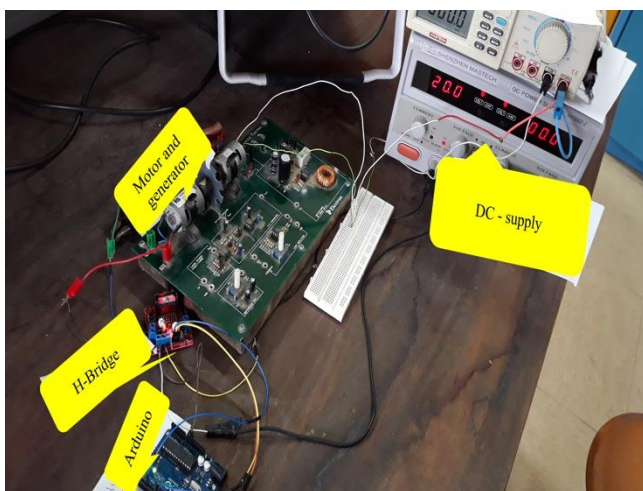


Fig. 2. The Hardware Connections for open loop system

generator, the module number of the board is B3510 – G- DC motor speed control, H-bridge, Arduino, voltmeter and power supply.

III. Experimental results of open loop system

In this section, the practical results of connecting the open loop system is showcased. The PWM signal generated from Arduino is sent to H-bridge, which a 20 V power supply was applied to H-bridge, to drive the motor and the output voltage is measured by multimeter.

It was noticed that any increase at the input voltage (V_{in}) by variation of modulation technique pulses, leads to an increase in the tachogenerator voltage (V_g), thus, the speed increases as can be seen in table I. Also, around 3 V is required for the motor to begin running due to the internal friction. With the fluctuations occurred and with the increase in the tachogenerator voltage accompanied with the increase in the speed of the motor, we conclude the open loop system is with high error.

Table I. RELATIONSHIP BETWEEN INPUT VOLTAGE AND GENERATOR VOLTAGE

Speed (0-255 rpm)	V_{in} (0-20V)	V_g (volts)
255	20	13.18
230	18	12.3
217	16	11.5
204	15	10.97
191	13	10
166	11	8.7
140	10	7.8
102	8	6
64	5	1.83
51	4	0.8
38	3	0
13	1	0

IV. Closed loop system description

A. Closed Loop DC Motor Speed Controller system

In this section, the feedback was applied to the system to acquire the closed loop system, which through the voltage of generator is returned to the Arduino, however the Arduino does not receive a voltage greater than the 5 V and the voltage of generator may be greater than this value, a voltage divider was used to protect the Arduino.

Fig. 3. represents the Arduino protection circuit. The voltage at the power supply equals to 20 V, the feedback to Arduino at pin A0 must not exceed 5 V, therefore, this circuit was used. It composes of two resistances connected in series, one of which is connected to Arduino so that the voltage is divided into two parts, and to assure the level of obtained voltage is less than 5 V the feedback was connected the on terminal of the smallest value resistance, the 5 K Ω resistance. Then, by

using the inverted voltage divider the real value of the voltage was obtained.

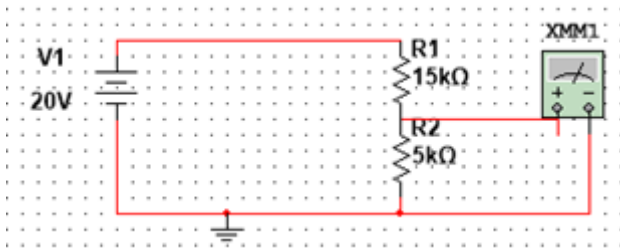


Fig. 3. Arduino protection circuit

The calculation of the aforementioned procedure is as follows:

Let v_1 is the voltage obtained from the terminals of 5kΩ resistance, where the supply voltage, v , is equal to 20 V

By using voltage divider:

$$v_1 = v * \frac{5k \Omega}{20k \Omega} , \quad (1)$$

$$\text{Then } v_1 = 5v \quad (2)$$

Here, it was observed that the maximum possible voltage value to be connected to the Arduino will not be higher than 5 volts and so that Arduino is protected from high voltages.

Then the lost value was offset in the code so the inverse of the voltage divider is applied, from equation (1)

$$\text{Then } v = v_1 * \frac{20k \Omega}{5k \Omega} \quad (3)$$

$$\text{Therefore } v = 4 * v_1 \quad (4)$$

The parameter's pins, the inputs and outputs were defined, then the direction of the DC motor was identified and the analog reading was converted into digital with a value between (0-5 V). For instance, the Arduino board contains six channels for converting similar signals into 10-bit digital, since the value of the voltage will be between 0 and 5 V it will be converted to a corresponding integer value between 0 and 1023. Thus, the accuracy of the operations of reading the values is the similarity: 5 V / 1024 units, then the voltage multiplied by 4, the factor that was obtained by equation (4).

H. The Hardware Connections

The hardware utilized in this research is the Arduino Uno, two motors, H-Bridge, variable resistor, and DC supply as illustrated in Fig. 4.

V. Expermental results of closed loop system

The main purpose of utilizing the PWM is to control the value of the input voltage, thus controlling the speed of the motor to convert to real value of V_{in} by the following equation

$$v_{in} = \frac{x}{255} * 20 \quad (5)$$

The range of x , PWM values, is between 0 to 255, the maximum voltage value at the maximum value of pulse width modulation (255) at Duty Cycle one (100 %). As shown in the Fig. 5 and illustrated in Table II.

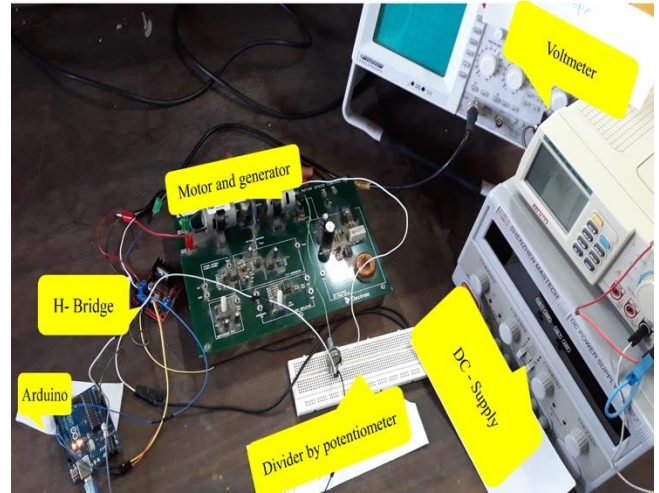


Fig. 4. The Hardware Connections for closed loop system

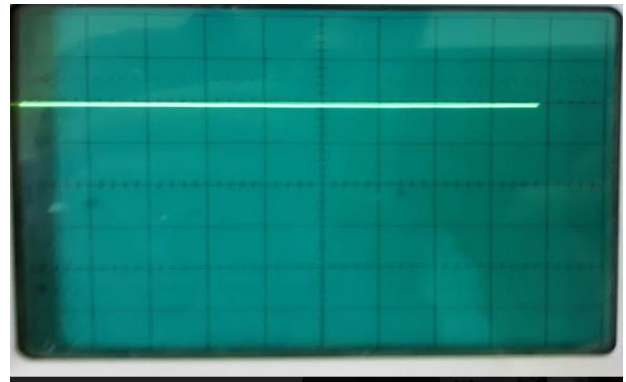


Fig. 5. The form of PWM while duty cycle equals to 100% (10 V/DIV)

Table II. AN EXPLANATION OF THE DIRECT RELATIONSHIP BETWEEN v_g AND v_{in} IN THE CLOSED LOOP SYSTEM

Speed (0-255 rpm)	V_{in} (0-20V)	V_g (volts)
255	20	13.72
200	15.7	11.4
166	13	9.85
140	11	9.15
102	8	6.47
77	6	2.33
64	5	2.04

The less the value of the duty cycle is, the less the value of PWM, leading to a decrease of the input voltage value. Fig. 6. shows the duty cycle when it equals to 50%.

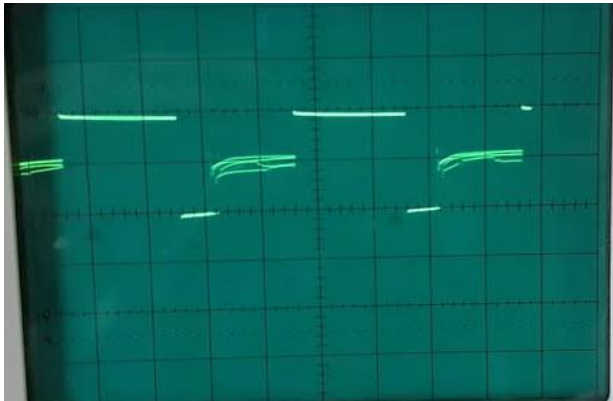


Fig. 6. The form of PWM while duty cycle equals to 50% (10 V/DIV)

As can be seen from the Fig. 7, the values of the generator voltage V_g through the connection of the closed loop system, therefore the system performance has been improved compared to the results obtained from the open loop system, since the distortion in the system has been eliminated.

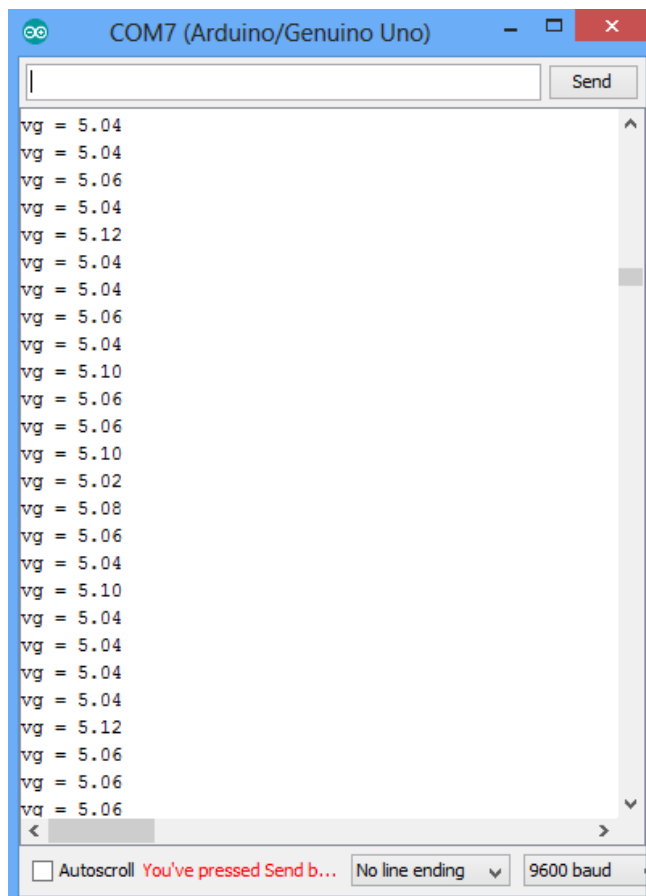


Fig. 7. The reading of generator output voltage V_g by Arduino.

In the closed loop system, the feedback voltage from the output was unstable, as a huge difference between the desired and actual voltage values occurred due to the error in the system, as illustrated in table II.

VI. Experimental description of closed loop system with PID controller

An PID controller is used to study the system response in terms of over-shoot, setting time and rise time and to try to set the PID parameters in order to get the most optimum possible response for the system to become stable.

A. The Experimental Set Up Explanation

First, we determined the direction of the motor by using the digital Write blocks as can be seen in the bottom side of the Fig. 8, as the switch of pin 7 is off and the switch of pin 8 is on, where they are the output of Arduino.

Then the unit step block at a constant equal to 10 volt was inserted to be the reference point, furthermore, a block for PID control was included to change the parameters of PID to control the system and make it stable. The signal then passes to a block that converts DC into pulses at output 3 passing from Arduino to H-Bridge to control the speed via the PWM. As was mentioned that the PWM is defined by the samples (from 0 to 255) and then the gain was inserted to convert it to the actual value.

Finally, a feedback block was added at the A0 port, and Summation was also included between the set Point and the Feedback voltage to make it a closed-loop system.

B. Experimental results

In the previous section, the closed-loop system without PID controller was applied, the system was unstable, as there was a large overshoot (which is indicated by the oscillation of the motor) and the steady state error that was also large.

However, through this section, a closed loop with the PID Controller to make the system stable was tested and the PID parameters (Proportional-Integral-Derivative) were adjusted to tune the best possible response. So, the proportional parameter k_p was increased to reduce rise time to the least, with a decrease in rise time, the overshoot increased. Moreover, by increasing the value of the integral parameter k_i to reduce the steady state error, while keeping k_p at a fixed value, the overshoot increases, and to overcome the high value of overshoot, the derivative parameter k_d was increased.

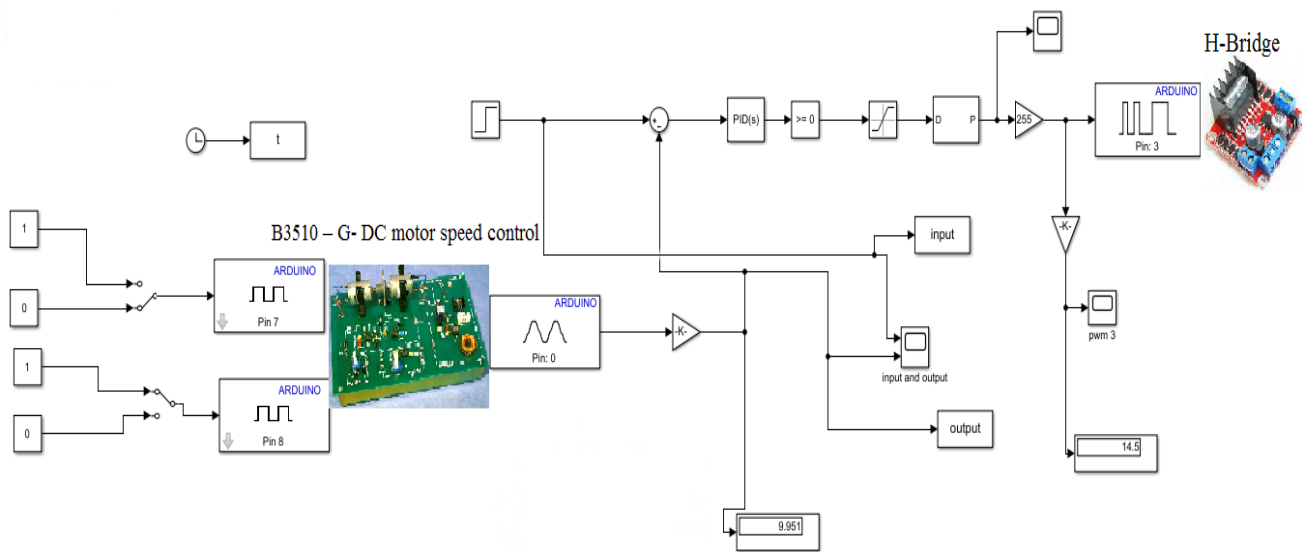


Fig. 8. Simulink of the system by adding a PID controller

Fig. 9. shows the response of the system after tuning the PID controller parameters manually, as it can be observed that there is no overshoot, with no steady state error, and a short rise time to reach the reference point, this result was acquired by setting the parameters with values k_p , k_i and k_d equal to 27, 110 and 12, respectively.

In this paper, the stability of the DC motor with generator load was applied using Arduino with Simulink to control the speed of the DC motor. It is reported that the stability of the DC motor using PID controller is more accurate. Furthermore, the system could be improved by using more than one intelligent control which was verified by simulation results. This approach will be used to analyze the stability of the time-delay DC motor in future work [7].

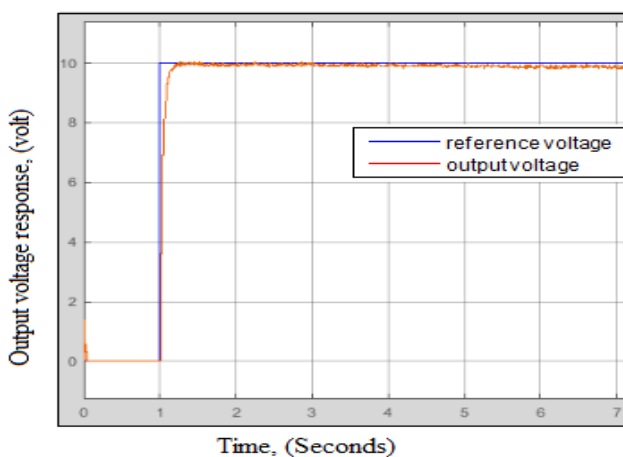


Fig. 9. The response of the system with the optimum value of k_p , k_i and k_d .

VII. Conclusions

In this paper a study was conducted to discuss the differences between the open loop system and the closed loop system of the B3510 – G- DC motor speed control module. Three cases of speed control were practically implemented; Open loop system, closed loop system and closed loop system using PID. The Arduino is used to control the module, the signal indicated by the Arduino was sent to the H-Bridge to drive the motor at a certain speed, according to the level of voltage resulted from PWM signal. The main purpose of utilizing the PWM is to control the value of the input voltage, thus controlling the speed of the motor to convert to real value of V_{in} . The speed is controlled through control the input voltage, using a PID controller with a closed loop system applied practically via utilizing Arduino and Matlab/ Simulink to generate the control signal. The results demonstrated that open loop system is more sensitive and the closed loop is more stable and less sensitive than open loop.

Acknowledgment

The Authors would like to acknowledge all staff members in the Electrical and Electronic Engineering department, at University of Benghazi who helped to realize and design this system with improved results.

References

- [1] Z. Has, A. H. Muslim and N. A. Mardiyah, "Adaptive-fuzzy-PID controller based disturbance observer for DC motor speed control," in 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2017
- [2] J. T. Teeter, M.-y. Chow and J. J. Brickley, "A novel fuzzy friction compensation approach to improve the performance of a DC motor control system," IEEE Transactions on Industrial Electronics, vol. 43, p. 113–120, 1996.
- [3] Y. A. Almatheel and A. Abdelrahman, "Speed control of DC motor using fuzzy logic controller," in 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), 2017.
- [4] Y. Guo and M. E. A. Mohamed, "Speed control of direct current motor using ANFIS based hybrid PID configuration controller," IEEE Access, vol. 8, p. 125638–125647, 2020.
- [5] A. Jaya, E. Purwanto, M. B. Fauziah, F. D. Murdianto, G. Prabowo and M. R. Rusli, "Design of PID-fuzzy for speed control of brushless DC motor in dynamic electric vehicle to improve steady-state performance," in 2017 International Electronics Symposium on Engineering Technology and Applications (IES-ETA), 2017.
- [6] B. Hekimoğlu, "Optimal tuning of fractional order PID controller for DC motor speed control via chaotic atom search optimization algorithm," IEEE Access, vol. 7, p. 38100–38114, 2019.
- [7] A. Khalil, O. Mohamed and J. Wang, "Wireless Networked Control Systems", 1-6, 2015.

Real-time Weather Station Monitoring System Through Wireless Sensor Network (WSN) Using off-the-shelf and Open-Source Components

¹Ezdihar Alwahhishee Ahmed, ²Mustafa KH. Aswad

¹ Electrical and electronic Engineering Dept. Faculty of Engineering

² Computer Engineering and IT Dept. Faculty of Engineering
Sabrahta Univeristy, Sabrahta, Libya

¹ Ezdihar.alw@gmail.com

² mustafaasawd@sabu.edu.ly

ABSTRACT

Keywords:

Arduino; DC motor;
PID; H-Bridge

Nowadays, DC machines, in particular motors, are widely employed in most applications. DC machines that convert the energy form from mechanical to electrical are called DC generators, whilst the DC machines that convert the electrical to mechanical are the motors. DC motors currently are the most commonly used due to their high torque at starting and easy maintenance and installation compared to their AC counterparts. Since the speed of the DC motor is crucial to the functioning of the motor, it is important to control it while varying the voltage. Therefore, this paper proposed controlling the speed through controlling input voltage, using a PID controller with a closed loop system applied practically via utilizing Arduino and Matlab/Simulink to generate the control signal. The results of connecting the Feedback system were compared to their counterparts while connecting the open loop system.

I. Introduction

Nowadays, Internet of Things (IoT) technologies are leveraged to serve people, provide faster and simpler services built around user needs, design to optimize resources, and make things more active. Thus, the Internet of Things is a combination of more introducing of emerging technology and more shifting out of human from many aspects of everyday life[1]. IoT improves almost everything. It provides number aspects which improves life such as an

intelligent cargo movement, cities become more digitalized and intelligent, companies will be able to track all of their products by means of radio frequency identification (RFID) tags.

Since, IoT technology emerge in our life, become very ease to automate, control and monitor any function of a house. Thing can be automate are: lighting, temperature control, multi-media, security, window and door operations, air quality.

This paper introduces a real time monitoring system utilizes seven major off-the-shelf components. The **Off-the-shelf** (OTS) component are: 1. The processing part composed of opensource Node MCU microcontroller with Wi-Fi, 2. The measurement part, consists of calibrated temperature sensors and gas detection sensors, 3. LED, for the weather forecasting, 4. RGB LED to monitor air quality level 5. The display part comprised of monochrome 0.96" 128X64 OLED and displayed via IoT platform, 6. Ventilation and air conditioning controllers using feedback from sensor-based, to control the room temperature. 7. The motion detection part, to control doors and windows.

Our goal is to build a robust system for sensors and data loggers that has the following features:

- Affordable price by utilizing low-cost and high-performance components compared to common commercial devices;
- Efficient implementation with an open-source hardware and software;
- System easy to build, program, and launch by modern expert engineer following detailed tutorials exit online.
- Data storage can be store on the storage media provide in the used device.
- Real-time data monitoring.
- Measurement accuracy similar to commonly used commercial devices.

The rest of the paper is structured as follows: an important aspects related to IoT are described in Section II. Section III presents our methodology. An Implementation Effort were described in Section VI. Finally, Section V presents the conclusion and Future Work.

II. IoT Aspects

The top three aspects that should be taken into account when an IoT system is built:

Privacy: with the fast increase in IoT claims a massive amount of data is produced where several security and privacy problems are observed. Even though the data encryption, hacker still can be able to extract information, by simply monitoring and exploring the traffic pattern[2].

Communication: Large number IOT devices involving a high amount of metadata transmission, e.g., the block headers in blockchains, conflicting the standard awareness that the IoT traffic has very small payloads[3]. The future of IoT will very much have to depend on decentralizing IoT networks.as result of connecting billion of billions and hundreds of billions of devices.

Compatibility: IoT is deploying many different technologies and standards. This will cause difficulties and require the deployment of extra hardware and software when connecting devices[4].

III. Methodology

We have selected off-the-shelf components based on power consumption, cost, and ease of usage to design and create the sensor node. The developed sensor node is equipped with a processing unit, several sensors, and a wireless

communication and mesh-networking module as shown in Fig. 1. The processing unit is the NodeMCU, an open-source microcontroller development board based on esp8266. Wireless communication is achieved by using the Wi-Fi. Different kind of sensors are used. The sensor nodes connect to the Internet through a Wi-Fi connection, sending its data to the database in the Cloud using a web API. The web API sends a warning text messages, or email about the situation of environmental conditions to the user.

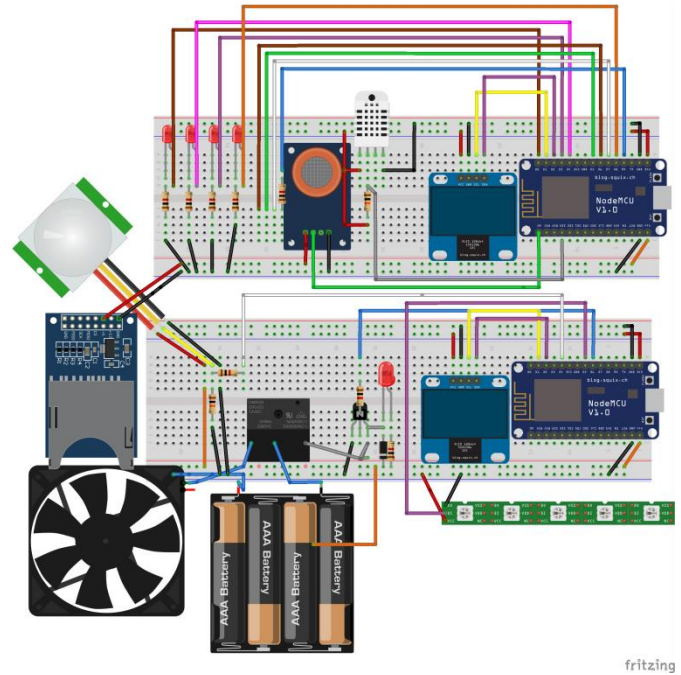




Fig. 1 The connection of NodeMCU with the DHT22, MQ9, PIR, Fan, LED, RGB LED, OLED, SD card, and the batteries (Open Source Hardware: Fritzing)

The following hardware, software, and application have been used in the experiment:

A. Hardware

1. NodeMCU-ESP8266 microcontroller with an integrated Wi-Fi with the following characteristics: low-cost price, small physical size, and low power consumption with USB connection port.
2. The rest components and sensors are listed in Table 1.

TABLE 1 – Measured Parameters and Sensors

	Type of Sensor	Sensors	Figures
Sensors	Temperature and Humidity	DHT22	
	Gas detection		

	Type of Sensor	Sensors	Figures
Component		MQ9	
	PIR motion detection	PIR	
	RGB LED	2812B	
	LED	3mm	
	Relay	5V	
	OLED	0.96"	

- MQTT is an open protocol established and released by IBM. To ensure the consistency of message transmission, MQTT supports three levels of QoS[5]: communication feature, whether at the local level or on the Internet, the need to develop techniques for the process of exchanging data between these devices also appears, and this protocol is one of the most widespread and reliable Internet of things protocols, and the simplified design of this protocol has made a suitable solution. For controllers or devices with limited and unlimited resources both in the field of processing and storage capacity. This protocol offers benefits that reduce power consumption and bandwidth, which are two very important factors in the world of IoT devices.

B. Software and Application

Many open source IoT platforms are available[9]. A comparison between nine IoT open sources platforms is shown in Table 2. For this experiment choosing one of the open source IoT platform fits our requirements is a non-trivial task. In our project, we have chosen the ThingsBoard because it is a 100% open-source available for cloud service and local hardware with Apache licence and Cassandra database. This IoT platform is easy to configure for data collection, processing, visualization, and device management. It enables device connectivity via industry standard IoT protocols - MQTT, CoAP and HTTP and supports both cloud and on-premise deployments in real time. ThingsBoard comes with scalability, fault tolerance and performance features.

Since we need a fully integrated system containing sensing, storage, and analytics. The following additional components have also been used:

- DHT22 sensor, to collect temperature data, regarding the Heating, Ventilation, and Air Conditioning (HVAC) system as shown in Table 1. The DHT22 sensor can be set to a certain threshold limit for maintaining the temperature, whenever the values exceeds the chosen threshold limit it triggers the system to take necessary actions for regulating the temperature.
- 4 LED, to monitor weather forecasting based on open weather map API. As shown in Figure (2,3) this device shows the weather forecast based on a change in the LED colour.

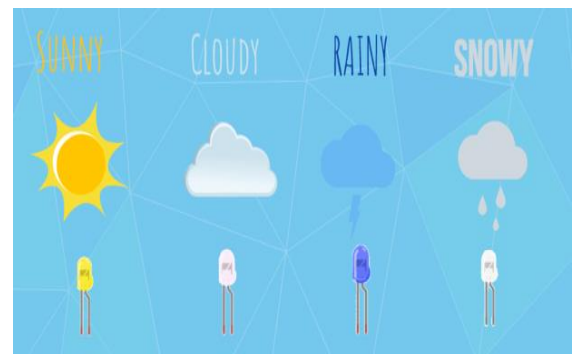


Fig 2 LED indications: Sunny, Cloudy, Rainy and Snow

- RGB LED, to monitor air quality levels as shown in Figure(2). The data are measured by MQ9 sensor, if exceeds the configured value range, the web app. sends an alert to the user, to take necessary action.

Table (2): a comparison between IoT open sources platforms (4)

IoT Software platform	Device Management	Protocols for data collection	DB
Kaa IOT	Yes	MQTT, CoAP, XMPP, TCP HTTP	MongoDB, Cassandra, Hadoop, Oracle, NoSql
SiteWhere	Yes	MQTT, AMQP, Stomp, WebSocket, and direct socket connection	MongoDB, Hbase, InfluxDB
Thing Speak	No	HTTP	MYSQL
DiveHive	Unknown	Rest API, WebSocket, MQTT	PostgreSQL, SAP HanaDB
Zetta	NO	HTTP	Unknown
Distributed Service Architectures (DSA)	NO	HTTP	ETSDB, Embedded Time SERIES
ThingsBords.io	Yes	MQTT, CoAP, and http	Cassandra
Thinger.io	Yes	MQTT, CoAP, and http	MongoDB
WSO2	Yes	HTTP, WSO2, ESB, MQTT	Oracle, PostgreSQL, mySQL

Know Before You Go! Air Quality Index		
Category	Index Value	Level of Health Concerns
Green	0-50	Good
Yellow	51-100	Moderate
Orange	101-150	Unhealthy for sensitive groups
Red	151-200	Unhealthy
Purple	201-300	Very unhealthy
Maroon	301-500	Hazardous

Fig.3 Air Quality Monitoring Shows Pollution level on OLED and RGB LED

IV. Implementation Effort

All parts were connected together to produce the transmitter part as showed in Figure 1. The transmitter part consists of a NodeMCU which has the wireless transceiver module, a DHT22 sensor, MQ9 (gas detector) sensor, and PIR motion detector sensor. The transmitter reads the values (temperature the humidity, gas leakage, etc.) every second, and sends these values to the receiver throughout the wireless module. Data is sent every second to make sure the receiver receives data in case of a lost package.

The transmitter and the receiver codes are very simple. The code was written in C, and it was about 180 lines of code. Setting up the whole system does not require a lot of effort from modern expert engineer. The project was easily to develop and the cost did not exceed 200LD.

The ThingsBoard smart meter dashboard has been used to allocate the hosted sensors online and on live server to visualize the represented smart meter IoT data that has been used in the project as shown in Fig. 4. The platform provides production ready server infrastructure to connect the smart meter devices, collect, store and analyze metering monitoring data, and to share results of the analysis with the end-users as shown in Fig. 4.



Fig. 4 Graphs and data for temperature, Humidity and Gas Sensors in WSN Monitoring Application (Open Source IoT platform: Thingsboard.io)

We have configured a set of rule on ThingsBoard to configure and customize the system behavior. It provides the REST API reference for server side integration to upload timeseries key-value data. The LEDs work as weather indicator and represent the quality of air as shown in Fig. 2. The obtained weather information (temperature, Humidity and Gas

Sensors) can be easily displayed on a globe map with its location.

V. Conclusion and Future Work

A control system is designed to remotely control and monitor the environmental conditions using off-the-shelf and open-source components (the existing WiFi infrastructure). The system was easy to build with minimum effort compared to traditional control systems. An IoT environmental solution described in this paper can be used to measure temperature, humidity, air pollution, gas leakages, HVAC system remote control, and weather forecasting. Such an integrated system could be modified to expand its functionality by incorporating features such as controlling the garden water system, check rainfall levels, wind speed and direction. A GPS module could also be added to determine the location. Also, we can easily integrate it with voice control devices such as Alexa Echo or Google Home.

IoT technology plays an important role in delivering in a cost effective way, a powerful efficient and flexible solution for business and end users, such as smart homes, data centre rooms, greenhouses, etc.

References

- [1] Kramp, Thorsten, Rob Van Kranenburg, and Sebastian Lange. "Introduction to the Internet of Things." Enabling Things to Talk. Springer, Berlin, Heidelberg, 2013. 1-10.
- [2] Rahman, L. F., Ozcelebi, T., & Lukkien, J. J. (2016, August). Choosing your IoT programming framework: Architectural aspects. In 2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud) (pp. 293-300). IEEE.
- [3] Danzi, P., Kalor, A. E., Sorensen, R. B., Hagelskjær, A. K., Nguyen, L. D., Stefanovic, C., & Popovski, P. (2020). Communication aspects of the integration of wireless iot devices with distributed ledger technology. IEEE Network, 34(1), 47-53.
- [4] Banafa, Ahmed. "Three major challenges facing iot." IEEE Internet of things (2017).
- [5] Lee, Shinho, et al. "Correlation analysis of MQTT loss and delay according to QoS level." The International Conference on Information Networking 2013 (ICOIN). IEEE, 2013.
- [6] Kodali, R. K., & Mandal, S. (2016). IoT based weather station. Dans 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 16-17 Dec. 2016 (pp. 680-683). IEEE. doi: 10.1109/ICCICCT.2016.7988038. Repéré à <http://dx.doi.org/10.1109/ICCICCT.2016.7988038>
- [7] Prinyakupt, J., & Yootho, T. (2016). Multichannel temperature monitor on IoT. Dans 2016 9th Biomedical Engineering International Conference (BMEiCON), 7-9 Dec. 2016 (pp. 4 pp.). IEEE. doi: 10.1109/BMEiCON.2016.7859623. Repéré à <http://dx.doi.org/10.1109/BMEiCON.2016.7859623>
- [8] Development of ambient environmental monitoring system through wireless sensor network (WSN) using NodeMCU and 'WSN monitoring'. Dans 6th Mediterranean Conference on Embedded Computing, MECO 2017, June 11, 2017 - June 15, 2017. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/MECO.2017.7977235. Repéré à <http://dx.doi.org/10.1109/MECO.2017.7977235>.
- [9] 9 Best & Top Open source IoT Platforms To Develop the IOT Projects, retrieved from <https://www.how2shout.com/tools/best-open-source-iot-platforms-develop-iot-projects.html> to Do (active Hd)

Applying Multiple Deep Learning Models for Antipersonal Landmines Recognition

¹ Abdelhamid Elwaer, ² Marwa Solla, ³ Fatima Ben Lashiher,
⁴ Hassan Ali Hassan Ebrahim, ⁵ Hala Shaari, ⁶ Rudwan A. Husain
^{1,2,3,4,5,6} Software Engineering Department, University of Tripoli, Tripoli, Libya
¹ ab.elwaer@uot.edu.ly, ² m.solla@uot.edu.ly, ³ f.benlashihar@uot.edu.ly,
⁴ H.Ebrahim@uot.edu.ly, ⁵ h.shaari@uot.edu.ly, ⁶ r.husain@uot.edu.ly

ABSTRACT

Keywords:

Arduino; DC motor;
PID; H-Bridge

Antipersonnel landmines represent a constant threat for the population of many countries. The huge amount of human lives lost due to this phenomenon has been a strong motivation for this research. Since, Deep Learning (DL) has already shown to be very effective in object detection, image classification, face recognition and other computer vision activities. This paper focuses on the problem of a landmine recognition in order to identify its type based on shape features. This research work consists of several stages: gathering a new dataset of Anti-Personnel Mines (APMs) images for training and testing purposes, employing several augmentation strategies to boost the diversity of training data, applying four different Convolutional Neural Network (CNN) models namely VGG, ResNet, MiniGoogleNet and MobileNet, and evaluating their performances on APMs recognition. In conclusion, results indicate that MiniGoogleNet exceed all of other three models in recognizing APMs with the highest accuracy rate of 97%.

I. Introduction

A landmine is an explosive object deployed to destroy or disable enemies. Landmines are implanted during war times hidden under or on the surface of ground and may remain undetected. The existence of explosive remnants of war (ERW), and specifically antipersonnel landmines, are considered as a serious threat for civilians as well as militaries around the World [1]. Many countries have suffered from mining for the last seven decades [2]. According to the UN Mine Action Service (UNMAS) regarding Libya, it has been reported that 108 were killed and another 161 were injured in 135 accidents since March 2011 [3]. The UN Department of Humanitarian Affairs (UNDHA) set a group of strict legislations that determine the strategy of

civil area demining. Never-the-less, there is no accurate estimation of these areas which were trapped by landmines. In fact, in order to consider an area free-mining, 99.6% of remnants of war must be safely eliminated. Generally, landmines detection, localization, and recognition are the steps followed by most landmine clearness systems. In reality, most of these systems depend on metal detectors and on deminer experience. Thus, with the purpose of protecting the lives of civilians and militaries, and facilitating demining operations, this paper focuses on the recognition problem which is essential to determine whether an object is an APM or not, and specify its model type [4].

Deep learning is part of Machine Learning (ML) that deals with considerable quantities of data. Recently, the technology

of DL has achieved great successes in object detection, image classification, face identification and other computer vision tasks. In fact, many supervised and unsupervised learning tasks were solved using multiple level model architectures [5] [6]. DL uses multiple layers represented in the CNN to generate computational models. The importance of classification derived from the variety of its uses in practical applications such as recognizing images of skin cancer, application of high-resolution imagery to identify disasters like floods, volcanoes, and droughts. Such use of DL will help noting the impacts and caused damages [7] [8].

This paper reports on a study applying four different architectures of deep learning modules for anti-personnel landmines recognition. The authors' motivation for using CNN comes from recent studies [9] [10] indicated that these DL architectures achieve a terrific success for several computer vision tasks such as image segmentation and identification. An important aim of this ongoing work is to find the most suitable model that allows landmine recognition with respectable accuracy. VGG, ResNet, MiniGoogleNet and MobileNet architectures, consisting of 16, 50, 75 and 28 layers respectively are employed for several experiments. A new dataset of landmine images, consists of 855 images, is used. This research consists of several stages: data collection, data augmentation, training of deep learning models, selection, and comparison. The paper has is structured with eight sections. Section I gives an introduction to the presented topic. That will be followed by section II which gives brief background on mines, deep learning and image recognition. Section III presents an overview of related work. Then, a description of the methodology is reported in section IV. The detailed description of the experiments setup and dataset is reported in Section V. Where the results are described in Section VI, Section VII evaluate the results. The final section VII presents the conclusions as well as the future work.

II. Background

A. Landmines

Mines came from the Latin word "mina" which means vein of ore which is an expression for digging a hole in the ground. Military engineers later used the term in relation with their job of digging landmines in the ground during military operations. At first, mines were applied to destroy goals were placed on earth, but afterwards, these mines were filled with explosive like gunpowder or black powder to cause greater damaged. Mines only began to appear on a large scale in 1918 in which mines against tanks first produced. They could be found along roads, in fields and forest, near wells and river bank causing serious economic problem for the countries. It was estimated that for every 5000 mines that are removed, one person is killed and two persons are injured. A variety of methods are used for the detection of landmines: Metal detector methods, Biological methods and Mechanical methods. With reference to the metal detector methods, sensors are used to measure the disturbances of the released electromagnetic field of metal objects that buried in the ground. Therefore, this method could not be used for plastic landmines. Whilst, the biological method depends on trained dogs, trained bees or

bacteria to detect mines covering large areas in short time. The mechanical method relies on clearing paths only using mechanisms such as remote-controlled clearness. Never the less, these methods are considered to be slow, expensive and dangerous [11]. Generally, landmines can be classified in to two types, APM and Anti-Tank mines (ATM).

B. Deep learning

Recently, artificial intelligence (AI) techniques such as ML and DL is being the focus of numerous interesting articles. Whilst AI was born in 1950, questioning computer scientists whether computers should think; machine learning came up with more specific question of whether a computer could accomplish a task without human inter. In order to obtain answers, set of rules and data should be applied by the machine learning. At first, producing rules using data and answers. Then these roles should be applied on new sort of data in order to produce answers for new questions. On the whole, ML systems are to be trained with expels linked to the area of a study with the intention of gaining rules to automate similar studies [13]

Currently, Deep Learning is the tendency in analyzing big data, and in learning systems. As an important area of machine learning that accomplishes classification tasks straight from images, video, texts or sounds, DL is described as a major enhancement of the Artificial Neural Networks (ANN). DL model architectures are consisting of a considerable number of neuron layers that give a superior level of abstraction and progression to the target data. Therefore, deep learning is count as the key automatic learning technique in computer vision and image processing fields [14]. Some of the advanced models of deep learning include AlexNet, VGG net, GoogleNet, ResNet from, and YOLO (You Only Look Once) [15]. CNN plays an important role mostly in analyzing visual data items. Although CNN was born in 1980s, it became a breakthrough in 2000 due to acceptance of GPU [16]. CNNs contain higher number of layers compared to the conventional NN, where each layer is trained for identifying certain features in an image. CNNs excel in maintaining spatial relationships with filtered input images. Such relationships are vital for realizing the variations between input image segments. Practically, filters are used in every training image with different resolutions in order to produce output that serves as input to the adjacent layer. At first, the filters produce simple features identifying edges, corners and lines. Afterwards, they generate more complex features that recognize the object [17]. A CNN architecture consists of a number of layers, which include a convolutional layer, a pooling layer, a ReLU layer, a fully connected layer, and a loss layer. Where the main layer is considered to be the Convolutional layer, that is due to the fact that it composed of kernel filters. The filters are to identify target image features, and the pooling layer is used to decrease the number of parameters in the NN. The ReLU layer represents the activation function that is used to set any negative weight values to the zero. The fully connected layer smoothens the output features. Finally, Loss layer illustrates

the differences between the predicted labels and the actual values [13].

III. Related Work

For demining purposes, a number of methods have been designed and developed. Conditions such as the type of the buried landmine, the substance of explosion and the nature of soil, determines the used method. In general, most of these detection methods are made of three main components; for capturing signs, a sensor is found; an image processor for handling the obtained signals; and for landmine recognition, a component of decision making is used [18]. Several studies focused on landmine detection rather than recognition. The authors of [19] proved that the standard deep learning techniques can be applied for automating image recognition of landmines. Furthermore, the study [20] proposed and assessed algorithms with the purpose of tackling the problem of detection and identification of landmines using GPRs. Genc and Akar [21] classified the techniques for buried target detection into four categories: shape-based, physics-based, and image-based techniques, and convolutional neural networks (CNNs). While the shape-based techniques are characterized by their ease, they are based on the hyperbola-like shape in B-scan image of ground penetrating radar (GPR) data. Hough transform [22] and alternative fitting techniques [23] are the most recognized shape-based object detection algorithms in GPR data. While Physics-based techniques attempt to assess the core properties of the buried object using its shape and size [24]. However, this technique is likely to fail as a result of the unidentified environment that the object is buried in [25]. Another well-known GPR-based landmine identification method is Image-based technique where features are extracted from B-scan image of GPR data in order to determine whether the buried object is threat or not [26-30]. Yet computational complexity is an obstacle [31][32]. While the three previous techniques rely on the separation between feature extraction and classification phases, the literature review shows that CNNs can deal with both phases cooperatively. As a matter of fact, landmine detection using CNNs in GPR B-scan data has been explored in several prior studies. CNNs results depends on the availability of a sufficient amount of data such as different landmines in various environments. However, as a large amount of existing GPR data of landmines is owned by military, it is hard to collect sufficient data to train CNNs [33-35].

IV. Methodology

A. Dataset and Preprocessing

A set of 855 samples of Landmines were obtained. The original dataset was collected for training and testing purposes, and it was gathered using one of the following three ways: Firstly, downloading images manually from landmines data websites such as CAT UXO [39]. This way consumed a lot of time simply due to the amount of human work involved. Secondly, using an automated python script works to download landmine images from the internet, which helped gathering images faster and with less effort. Finally, a few images were captured from real field in Libya using a

mobile phone. The dataset is divided into 7 classes of Landmines images which are named according to the shape as, Butterfly Shape, Cylindrical Fan Around, Cylindrical Flat, Cylindrical Upper Bump, Cylindrical Upper Fan, Screw Shape, and Rectangle Shape. Some images are given in Figure 1. Additionally, the size of a training dataset was expanded using data augmentation. Therefore, a new dataset of 7000 samples were generated of augmented images after using data augmentation technique.



Fig 1. Sample examples for each class of Landmines.

The following table summarizes the dataset. approximately 80% of the images in the dataset were in the training set, and 20% were in the validation set.

TABLE I. DATASET DETAILS

No of Class	Classes	Number of images
1	ButterflyShape	1000
2	CylindricalFanAround	1000
3	CylindricalFlat	1000
4	CylindricalUpperBump	1000
5	CylindricalUpperFan	1000
6	ScrewShape1	1000
7	RectangleShape	1000
Total		7000

B. Data Augmentation

In order to build effective Deep Learning models, it is important to keep the validation error as low as possible while keeping the training error to a minimal. The effectiveness of this strategy may be proven by using data augmentation. Image data augmentation includes several techniques used for increasing the size of training datasets and add variety to datasets in order to build more accurate deep learning models. However, modified versions of original images were added into the dataset [40]. It is considered as a powerful technique for improving the accuracy of modern image classifiers [41].

In many machine learning tasks, Data augmentation is a commonly used technique. A variety of augmentation strategies have been proposed and shown to capture

important characteristics of natural images. The augmentation strategies can generate variants of the images, which can increase the fit models' capacity to generalize their knowledge to new images.[42]. One of the more effective data augmentations strategies is the traditional transformations, include flipping, Rotations, Zooming, and shifting. To perform the basic transformation, only affine transformations are used [43]. Each input image was output in two distinct forms; an exact copy shifted (width/height), zoomed out, rotated, flipped(horizontal), sheared, or rescaled. Eight augmentation strategies were utilized to generate new training sets. The following Table 2, illustrates the data augmentation techniques used in this experiment for randomly generated images.

TABLE II. AUGMENTATION TECHNIQUES

Technique	Description	Value
Rotation	Rotational augmentation can be performed by turning the image clockwise (between 1° and 359°) or counterclockwise (between 359° and 1°). Rotate the image between 0 and 40 degrees for the value of 40.	40
Rescale	In scaling or resizing, the image is resized to the given size.	64x64
Shear	Shear the image with increasing rate magnitude along the horizontal (vertical) axis. The value 0.2 meant shearing the image by 0 to 20 degrees.	0.2
Zoom	Zooming is for randomly zooming inside pictures. The value 0.2 means zoom-in and zoom-out by 20%	0.2
Flipping	It is for randomly flipping half of the images. Randomly flip inputs horizontally.	True
Width shift	Shift the image by 20 % along the X-axis.	0.2
Height shift	Shift the image by 20 percentage along the Y-axis.	0.2
Fill mode	Filling the area that was left over of shifting with the nearest pixel and stretching it.	Nearest

To illustrate the effect of augmentation strategies that were used for the purpose of this study, Figure 2, presents a set of the training images before and after augmentation.



Fig 2. Examples of original image vs augmented images

V. Experiments

In the experiments, the authors adopted four deep neural network models called VGG, ResNet, Minigooglenet, and MobileNet. These models were trained with a high-performance computing (HPC) unit, which has the following specifications: AMD Ryzen 2600 CPU, 16 GB RAM, and Nvidia GTX 980TI. Data processing and training was carried out using Keras library [44]. This API was built using Python, and runs on TensorFlow. [45]

A. VGGNet Model

The VGG16 CNN architecture, as its name suggests, is made up of 16 layers of neurons. It was proposed in 2014 by Simonyan and Zisserman of the University of Oxford.[46]. The majority of this network is made up of convolution and dropout layers. Normalization, concatenations, and shortcuts were excluded. This model achieves an accuracy of 89.88 percent on top-5 test accuracy in ImageNet, which has over 14 million images belonging to 1000 classes [46]. The VGGNet was very straightforward. It contained only three * three convolutional receptive field layers; it was capable of differentiating fundamental directions only. Additionally, it contained three completely connected levels and two pooling layers with a capacity of two.

B. ResNet Model

A residual neural network (ResNet) was launched in 2015 by He et al. [47], ResNet, or Residual Network trains Convolutional Neural Networks to previously unthinkable depths by utilizing what is known as a residual module. and can gain accuracy from this increased depth. With Resnet, networks with more than 100 layers may be effectively trained on the difficult ImageNet dataset. ResNets divide convolutional layers into Residual Blocks, and each block is given a Residual Connection by passing the corresponding block. the residual block's output is merged by summation with the original input that was forwarded by the residual connection. Eight times as deep as VGG nets, residual nets with a depth of up to 152 layers were tested on the ImageNet dataset. ResNet's depth was 20 times that of AlexNet. In our experiments a ResNet of 50 layer depth was adopted

C. MiniGoogleNet Model

MiniGoogleNet, as suggested by its name, it is a smaller version of GoogleNet model. It has simpler architecture complexity, and it was explored as logo detector to decrease the number of model parameters and the workload associated with training. The MiniGoogleNet architecture is built on three main modules: First, there is a Conv Module that is responsible for convolution, followed by Batch Normalization and ultimately, activation. Second, Inception Module which is like a two-branched mini-inception module, one is module that learns 1×1 filters, as well as the other one is also a Conv Module that learns 3×3 filters. Third, Downsample Module, which is responsible for decreasing an input volume's spatial dimensions. MiniGoogleNet was tested using three open datasets. The model achieves 90% accuracy on CIFAR-10 [48].

D. MobileNet Model

MobileNet model is built on a simplified design that It utilizes depthwise separable convolutions to construct lightweight deep neural networks. The model introduces 2 global hyper-parameters which make an effective trade-off between latency and accuracy. Depthwise Separable Convolution is a core layer of MobileNet [49], Which is a type of factorized convolution in which a conventional convolution is factorized into a depthwise and pointwise convolution. In the first convolution, each input channel receives a single filter. The second convolution then uses one * one convolution to combine the depthwise convolution outputs [50]. the first layer, which constitutes a complete convolution, the MobileNet has 28 layers. In ImageNet, the model obtains an accuracy of 0.895 percent for top-5 tests [49].

To summarize, the following Table 3, indicates a comparison between VGG, Resnet, MiniGoogleNet and MobileNet models. ResNet exhibits one of the highest performance accuracies in the ILSVRC 2012 dataset, outperforming VGG and GoogleNet.

TABLE III. COMPARISON MODELS

Models				
	VGG	ResNet	MiniGoogleNet	MobileNet
Published	2014	2015	2016	2017
Size	528MB	98MB	20MB	25MB
Top-5 accuracy	0.901	0.921	0.900	0.895
Parameters	138,357,544	25,636,712	1,662,295	4,253,864
Layers	16	50	75	28
Default input size	224x224	224x224	224x224	224x224

These pretrained models can be evaluating using various metrics, such as Recall, F Score, precision and accuracy. However, the final results are based on accuracy. Precision is a term that refers to the proportion of Predicted Positive situations that are in fact Real Positives. It could be defined as the proportion of correct positive results to the proportion of anticipated positive results by the classifier.[51]. Conversely, Recall or sensitivity refers to the proportion of True Positive cases that are successfully predicted as True Positive. By the +P (Predicted Positive) rule, the Coverage of Genuine Positive instances is measured. to sum, Recall is defined as the proportion of correct positive outcomes to the total number of relevant samples. while the F1-score is the arithmetic mean of Precision and Recall. Accuracy is defined as the ratio of correct predictions to total input samples. Commonly, accuracy is the most applied metric to evaluate the model performance [51].

In our experiments, The Four models were setup to be experimented on the landmine's dataset. When the data and labels are loaded to the model, training and testing split are performed, 75 percentage of the data is used for training, while the remaining 25 percentage is used for testing. Our images were preprocessed before they were fed to the models; every image in our dataset was resized to 64×64

pixels. Once the image is resized, it is scaled to the range [0,1]. In the all experiments, The cross-entropy [52] is applied as our loss function, and as our optimizer, stochastic gradient descent is used with a learning rate of 0.005. All models were trained for 150 epochs using mini-batch sizes of 32.

VI. Results

Examining this figure 3, we can see that MiniGoogleNet is obtaining 99% training accuracy. Furthermore, looking at our accuracy and loss plot over time in the figure demonstrates that our network is behaving quite well, after only 20 epochs the network is already reaching $\approx 99\%$ classification accuracy. Loss on both the training and validation data continues to fall with only a handful of minor "spikes" due to our learning rate staying constant and not decaying. At the end of the 40th epoch, we are reaching 99% training accuracy on our training set.

Training Loss and Accuracy on Landmines for seven class

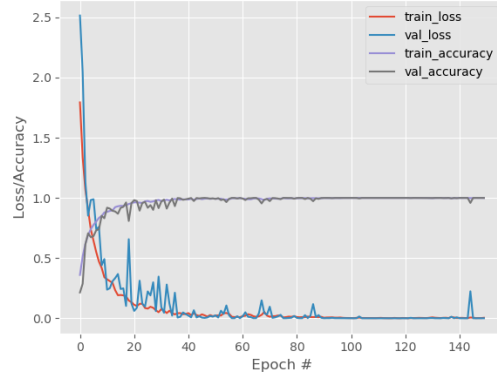


Fig3. MiniGoogleNet's result on Landmines dataset

The above plot shows a quintessential graph as the loss decreases each time the accuracy increases, moreover the training and validation loss and accuracy mimic each other indicating that our network is learning the underlying patterns without overfitting. In addition, the result illustrates in Figure 4 shows that after only 50 epochs the MobileNet model achieved an overall accuracy of $\approx 99\%$ training accuracy.

Training Loss and Accuracy on Landmines for seven class

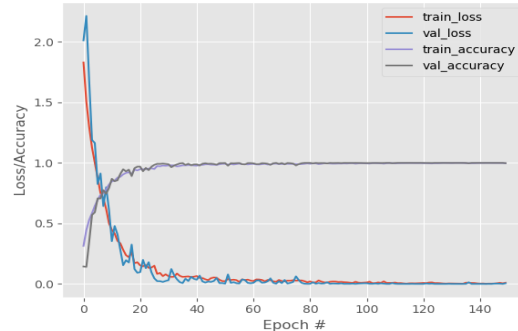


Fig4. MobileNet's result on Landmines dataset

As well as, the result illustrates in Figure 5 shows that after only 20 epochs the ResNet model achieved an overall accuracy of $\approx 99\%$ training accuracy.

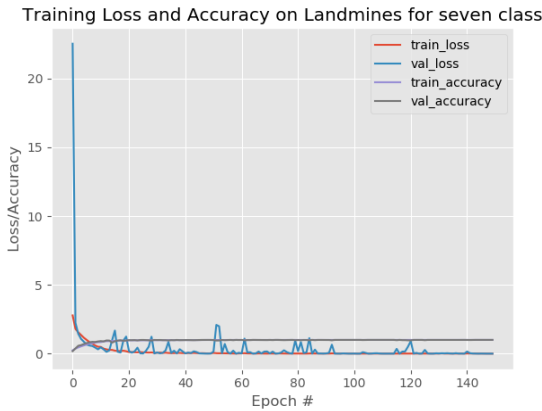


Fig5. ResNet's result on Landmines dataset

Furthermore, the result illustrates in Figure 6 shows that after only 60 epochs the VGG model achieved an overall accuracy of $\approx 99\%$ training accuracy.

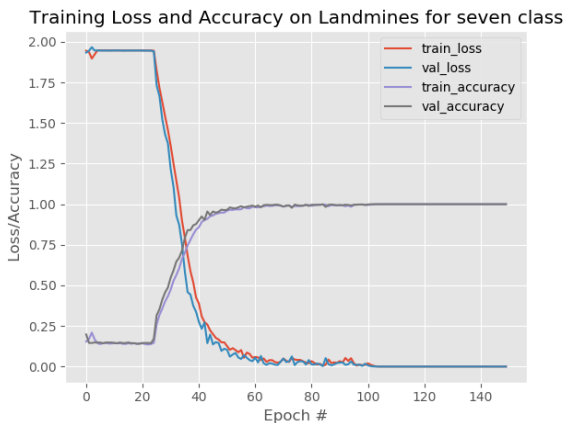


Fig6. VGG's result on Landmines dataset

The models' performance was determined by the ability to predict the correct landmines type on 1750 testing images. The best performing model is MiniGooglenet which achieves a mean F1 score of 0.9714, an overall accuracy of 97.0%. The experiments' results are tabulated below in table 4, it is apparent that all of the models showed a different performance-

TABLE V. COMPARISON BETWEEN RESNET, VGG, MOBILENET, AND MINIGOOGLENET

Performance Measures	VGG	ResNet	Minigooglenet	MobileNet
Precision	59.42	85.57	97.42	85.42
Recall	59.71	86.00	97.42	85.85
F-Score	59.42	85.71	97.14	85.42
Accuracy	60.00	86.00	97.00	86.00

Starting with the Precision, VGG had the lower results with 59.42%, followed by MobileNet with 85.42%, ResNet 50 with 85.57%, and the best Precision result with MiniGoogleNet, which obtained 97.42%. In the same way, the order of the Recall results was equal to the previous measure, where the lower percentage was obtained by VGG 59.71%, and the highest by MiniGoogleNet with 97.42%. On the other hand, F-score measurements, VGG achieved lower result with 59.42%, followed by MobileNet, ResNet, and MiniGoogleNet with 85.42%, 85.71%, and 97.14% respectively. Finally, in the accuracy metric VGG performance was poor with 60.0%, followed by MobileNet, ResNet which are showed a similar result with 86.0%, and highest result was obtained by MiniGoogleNet with 97.0%. Indeed, as is shown, the MiniGoogleNet implementation achieved the highest percentage.

VII. Discussion

Deep Learning provides a good opportunity to expand a research and the application of the basis for Landmine classification using digital images. The accurate models presented by deep learning are required to recognize the landmines. For this work, four models, namely VGG, Resnet, MiniGoogleNet, and MobileNet were trained. As well as, the dataset used provides seven classes of the landmines, which have a total of 6995 images. the dataset was extended using data augmentation techniques. The dataset was allocated 75% to training, and 25 percent for testing. As observed in the results section, All of CNN models were evaluated by various metrics: precision, Recall, F-Score, and accuracy, using the same epochs. Based on Table 4, which shows a general result performance, MiniGoogleNet achieved better results than the other model architectures. Both neural networks, ResNet and MobileNet were had Converging result. On the other hand, VGG shows lower performance. Additionally, although all models were trained using an equal number of epochs at 150, it can be noted that the training process can be stopped early without affecting the accuracy rate. However, based on the results of the experiment on models, the accuracy was starting to be stabilized at 40 epochs in Minigooglenet, MobileNet, and ResNet, as indicated in Figures 3, 4, and 5, respectively. Correspondingly, as shown in Figure 6, VGG model recorded that the accuracy stops keeping grow up at 60 epochs, which mean that the network accuracy starts saturating. In fact, this pointed that the training process can be safely stop at a smaller number of 150 in our experiments To sum up, this paper showed that deep learning models were feasible and beneficial in antipersonnel landmine recognition. In addition, it could be beneficial to create a landmine recognition mobile system that implements the best model for recognition of Antipersonnel Landmines, MiniGoogleNet is a good recommendation predictor for building a mobile application, so users with a little or no knowledge can use it to recognize of Landmines.

VIII. Conclusion

This paper demonstrated the usage for models of deep learning in a new application of landmines recognition. In recent years, deep learning technologies in image recognition has accomplished great successes. For the purposes of this paper, CNN four different models were used to verify the classes of the proposed dataset. A new APMs dataset was collected as part of this ongoing research. In order to expand the training dataset, traditional data augmentation methods - based on a mixture of affine image transformations are fast and simple to apply were adapted and have shown that data augmentation is the most often used technique for decreasing overfitting and increasing the dataset size. The CNN models MiniGooglenet, ResNet, MobileNet, and VGG were used to recognize APMs. The findings of the experiments reveal that three models achieve a good level of accuracy between 86% and 97%, on 1750 test images while last model achieves poor result. In fact, MiniGooglenet seems to be the valuable choice for this experiment with the best results on landmine dataset. To conclude, it has been demonstrated that convolutional neural networks provide excellent results in image recognition. For upcoming work, our future goal, therefore, is to improve our current landmine dataset by increasing the number and the diversify of its images is a long-term plan. Experiments on different models and comparing them are another objective. In the near future, the main goal is employing the chosen DL Model and the generated dataset to develop an APM recognition mobile application. That would be a great assistance for dominer in landmine fields.

Acknowledgments

The authors are the founding members of Data and Multimedia Lab at the Faculty of Information Technology in the University of Tripoli, <https://uot.edu.ly/it/dml/>. They are grateful to the support and contributions provided to them and their research lab.

References

- [1] A. Shire, U. Jawarkar and M.S. Chavhan, "Overview of Buried Mines Detection Using Radar Bullet", *International Journal of Engineering Sciences & Research Technology*, vol. 4, no. 2, pp. 250-254, 2015.
- [2] John, P. J. (2019). Review Paper on Object Detection using Deep Learning- Understanding different Algorithms and Models to Design Effective Object Detection Network. *International Journal for Research in Applied Science and Engineering Technology*, 7(3), 1684–1689. <https://doi.org/10.22214/ijraset.2019.3313>
- [3] Newsletter, F., & Africa, N. (2021). Libya struggles to clear explosive remnants of war. *The New Humanitarian*, 1–9. Retrieved from <https://www.thenewhumanitarian.org/analysis/2013/08/06/libya-struggles-clear-explosive-remnants-war/>
- [4] P. Bestagini, F. Lombardi, M. Lualdi, F. Picetti and S. Tubaro, "Landmine detection using autoencoders on multi-polarization GPR volumetric data", *arXiv:1810.01316*, 2018, [online] Available: <https://arxiv.org/abs/1810.01316>.
- [5] Hertel, L. (2014). Object Recognition Using Convolutional Neural Networks. Chapter (November). <https://doi.org/10.5772/intechopen.89726>
- [6] Fujiyoshi, H., Hirakawa, T., & Yamashita, T. (2019). Deep learning-based image recognition for autonomous driving. *IATSS Research*, 43(4), 244–252. <https://doi.org/10.1016/j.iatssr.2019.11.008>
- [7] Mustafid, A., Pamuji, M. M., & Helmiyah, S. (2020). A Comparative Study of Transfer Learning and Fine-Tuning Method on Deep Learning Models for Wayang Dataset
- [8] Menezes, R. S. T. d., Magalhaes, R. M., and Maia, H. (2019). Object recognition using convolutional neural networks. In *Artificial Neural Networks*. IntechOpen.
- [9] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European Conference on Computer Vision*. Springer, 2014, pp. 818–833.
- [10] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," *arXiv preprint arXiv:1312.6229*, 2013.
- [11] A. Shire, U. Jawarkar and M.S. Chavhan, "Overview of Buried Mines Detection Using Radar Bullet", *International Journal of Engineering Sciences & Research Technology*, vol. 4, no. 2, pp. 250-254, 2015.
- [12] R. Achkar, (2012). Implementation of a Vision System for a Landmine Detecting Robot Using Artificial Neural Network. *International Journal of Artificial Intelligence & Applications*, 3(5), 73–92. <https://doi.org/10.5121/ijai.2012.3507>
- [13] M. Shu., (2019). Deep learning for image classification on very small datasets using transfer learning. *Creative Components*, 14–21.
- [14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
- [15] M. Shu, "Deep learning for image classification on very small datasets using transfer learning," M.S. thesis, Iowa State Univ., Ames, IA, USA, 2019
- [16] C. McClanahan, "History and evolution of gpu architecture," *A Survey Paper*, p. 9, 2010.
- [17] H. Greenspan, B. Ginneken, and R. Summers, "Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1153-1159, 2016.
- [18] H. Kasban, O. Zahran, Sayed M. Elaraby, M. ElI Kordy and F. E. Abd El-Samie, "A Comparative Study of Landmine Detection Techniques", *Sens Imaging* 11:89–112, Springer, 2010
- [19] C. M. Ward et al "Deep learning for automatic ordnance recognition," in [Geospatial Informatics IX], 10992, 109920H, International Society for Optics and Photonics (2019).
- [20] K. H. Ko et al "GPR-based landmine detection and identification using multiple features," *Int. J. Antennas Propag.*, Vol. 2012, 1–7, 2012.
- [21] A. Genc and G. B. Akar, "Combination of physics-based and image-based features for landmine identification in ground penetrating radar data", *J. Appl. Remote Sens.*, vol. 13, no. 2, pp. 1, Apr. 2019.
- [22] G. Borgioli et al., "The detection of buried pipes from time-of-flight radar data," *IEEE Trans. Geosci. Remote Sens.* 46(8), 2254–2266 (2008).
- [23] H. Chen and A. G. Cohn, "Probabilistic robust hyperbola mixture model for interpreting ground penetrating radar data," in *Int. Joint Conf. Neural Networks (IJCNN)*, pp. 1–8 (2010)
- [24] R. Firoozabadi et al., "Subsurface sensing of buried objects under a randomly rough surface using scattered electromagnetic field data," *IEEE Trans. Geosci. Remote Sens.* 45(1), 104–117 (2007)
- [25] O. Lazaro-Mancilla and E. Gomez-Trevino, "Ground penetrating radar inversion in 1-D: an approach or the estimation of electrical conductivity, dielectric permittivity and magnetic permeability," *J. Appl. Geophys.* 43(2), 199–213 (2000).
- [26] T. Wang et al., "Frequency sub-band processing and feature analysis of forward-looking ground penetrating radar signals for landmine detection," *IEEE Trans. Geosci. Remote Sens.* 45(3), 718–729 (2007)
- [27] D. P. Williams, V. Myers, and M. S. Silvious, "Mine classification with imbalanced data," *IEEE Geosci. Remote Sens. Lett.* 6(3), 528–532 (2009).
- [28] E. Pasolli, F. Melgani, and M. Donelli, "Automatic analysis of GPR images: a pattern recognition approach," *IEEE Trans. Geosci. Remote Sens.* 47(7), 2206–2217 (2009).

- [29] J. S. Kobashigawa et al., "Classification of buried targets using ground penetrating radar: comparison between genetic programming and neural networks," IEEE Antennas Wirel. Propag. Lett. 10, 971–974 (2011).
- [30] A. B. Yoldemir and M. Sezgin, "A least squares approach to buried object detection using ground penetrating radar," IEEE Sens. J. 11(6), 1337–1341 (2011).
- [31] H. Frigui and P. Gader, "Detection and discrimination of landmines in ground penetrating radar based on edge histogram descriptors and a possibilistic k-nearest neighbor classifier," IEEE Trans. Fuzzy Syst. 17(1), 185–199 (2009).
- [32] D. Reichman, L. M. Collins, and J. M. Malof, "Improvements to the histogram of oriented gradient (HOG) pre-screener for buried threat detection in ground penetrating radar data," Proc. SPIE 10182, 101820U (2017).
- [33] S. Lameri et al., "Landmine detection from GPR data using convolutional neural networks," in Proc. 25th Eur. Signal Process. Conf EUSIPCO, pp. 508–512 (2017).
- [34] D. Reichman, L. M. Collins, and J. M. Malof, "Some good practices for applying convolutional neural networks to buried threat detection in ground penetrating radar," in Proc. IWAGPR, pp. 1–5 (2017).
- [35] L. E. Besaw and P. J. Stimac, "Deep convolutional neural networks for classifying GPR B-scans," Proc. SPIE 9454, 945413 (2015).
- [36] Hinge Ltd, Collective Awareness to Unexploded Ordnance (CAT-UXO), 2021, <https://cat-uxo.com>. [accessed 2019 Oct 1]
- [37] C. Shorten T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning". Journal of Big data. Pp.1-48(2019) 6:60. <https://doi.org/10.1186/s40537-019-0197-0>.
- [38] Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, and Quoc V. Le, "AutoAugment: Learning Augmentation Strategies From Data". Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 113-123, 2019.
- [39] Jason Brownlee, "deep learning for computer vision: image classification, object detection and face recognition in python". 563 pages. This guide online at: <http://machinelearningmastery.com> 2019.
- [40] Jason Wang , Luis Perez . "The Effectiveness of Data Augmentation in Image Classification using Deep Learning". (13-12-2017). <https://arxiv.org/abs/1712.04621>.
- [41] Aakash Nain, Sayak Paul, Margaret Maynard-Reid, Keras, <https://keras.io/api/preprocessing/image/#imagedatagenerator-class>. [accessed 2019 Oct 1]
- [42] Tensorflow, <https://www.tensorflow.org>. [accessed 2019 Nov 15]
- [43] Karen Simonyan, Andrew Zisserman., "Very Deep Convolutional Networks For Large-Scale Image Recognition". (10-04- 2015). <https://arxiv.org/abs/1409.1556v6>.
- [44] He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian , "Deep Residual Learning for Image Recognition". (10-12-2015). <https://arxiv.org/abs/1512.03385>.
- [45] Adrian Rosebrock, "Deep Learning for Computer Vision with Python: Starter Bundle". PyImageSearch, 2017, Pages 330.
- [46] Andrew G. Howard, Menglong Zhu, Bo Chen, and Dmitry Kalenichenko, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications". (17 Apr 2017). <https://arxiv.org/abs/1704.04861>.
- [47] Felix Altenberger , Claus Lenz." A Non-Technical Survey on Deep Convolutional Neural Network Architectures".(6 Mar 2018.). <https://arxiv.org/abs/1803.02129>.
- [48] David M W Powers. "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation". Technical Report SIE-07-001 December 2007. https://web.archive.org/web/20191114213255/https://www.flinders.edu.au/science_engineering/fms/School-CSEM/publications/tech_reps-research_artfcts/TRRA_2.
- [49] Katarzyna Janocha and Wojciech Marian Czarnecki, "On Loss Functions for Deep Neural Networks in Classification," Theoretical Foundations of Machine Learning (2017). <https://arxiv.org/abs/1702>.

Design and Implementation of a Mini-Automated Teller Machine Replica

Taha M. Abujrad, Mohamed M. Edardar

^{1,2} Department of Electrical and Electronic Engineering, University of Tripoli, Tripoli, Libya

¹ t.abujrad@uot.edu.ly, ² m.edardar@uot.edu.ly

ABSTRACT

Keywords:

ATM, ESP32, RFID, LCD, FSM

In this paper a small 5v physical model of the Automated Teller Machine based on the use of the ESP32 microcontroller is designed and implemented. This model provides most of the main services (withdrawal, transfer, deposit, balance enquire) implemented using knowledge obtained from various disciplines (control, digital, computer, networks, and mechanics). To perfectly simulate the ATM transactions, a virtual bank was created, the ATM can connect to this bank via the internet to fetch the needed information from the costumer's account and updating the relevant information when the transaction finishes. This virtual bank is saved as a database in a public domain named <https://atm.root.ly>. Furthermore, in this model RFID based cards were used, where each card carries a unique identifier (UID) rather than using magnetic tape cards. This replica contributes to the academic society as being a noncommercial test model used by researches studies that enhances the features of the existing ATMs. It also contributes as being a device that can be used by firms that want to start using the real ATM, which will give them an idea about the benefits of ATMs on one hand, and the precautions that should be taken when using them on the other hand.

I. Introduction

An automated teller machine (ATM) is an electronic banking outlet that allows customers to complete basic transactions without the aid of a branch representative or teller. To use an automatic teller machine, clients must have a plastic ATM card with a plastic smartcard with a chip or a magnetic stripe, which contains a unique card number and some security information about the client. The customer is identified by inserting plastic ATM card and entering a personal identification number (PIN). But due to several security challenges associated with these cards including frauds and thefts; recent studies [1] suggests using biometric indices (such as the iris, and the fingerprint) rather than plastic cards since they provide promising tools that eliminates all the security problems. In either situation, the ATM sends quires to the customer's database, requesting for his balance, information about last transaction, taxes ... etc. After the transaction successfully completes the ATM updates the customer's relevant

information by sending the appropriate queries to the server. [2]

The conceptual model of an ATM is often represented in Finite state machine format (FSM). This FSM model adopts a set of states (ATM services in this case) and state transitions (ATM system functions).[3]

Figure-1, shows the generalized FSM model of this ATM, where Service X represents any service this ATM offers, the main four services are listed in Table-1, along with their requirements and their end results.

There exist previous attempts for the design of an ATM, as in [4, 5] and other attempt [6] for the design to a vending machine (which is similar to the ATM in its automated behavior) that worth mentioning, unfortunately none of them implemented the hardware part of the machine, they seldom focused on the software part regarding the two factor authentication method, and the user interface to the

machine. Hence this paper is innovating in its mechanical design idea.

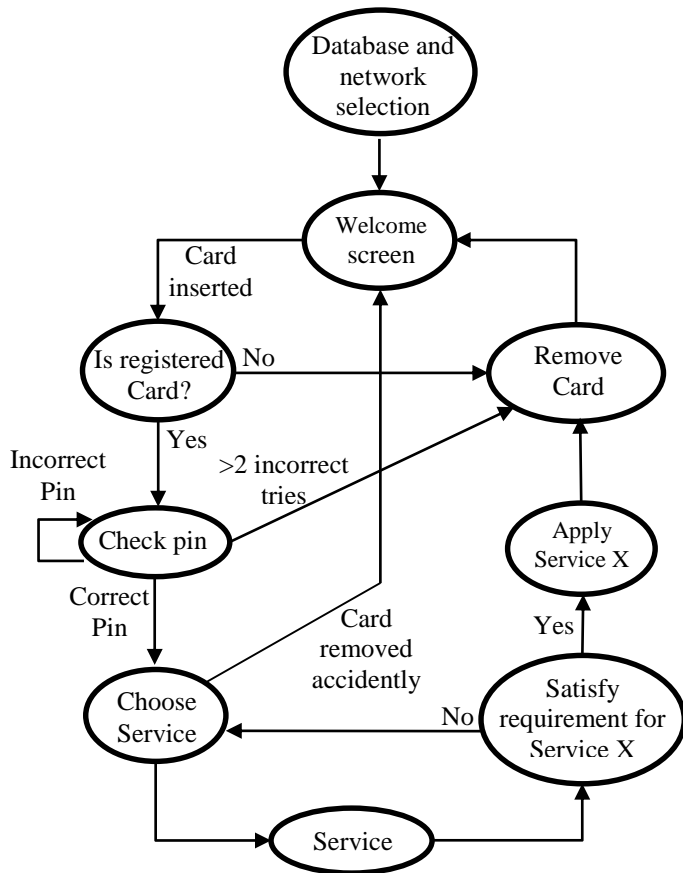


Fig. 1 Simplified Finite State Machine Model of the ATM

Table 1 ATM services requirements and applications

SERVICE	REQUIRMENTS	END RESULTS
Withdraw	Enough Balance, enough notes, not exceeding withdrawing limit	Distribute the entered/selected amount using the least available denomination principle
Transfer Money	Enough Balance, the recipient customer knows his pin number	The entered amount is transferred from the doner to the recipient
Deposit	The customer be the manager of the ATM	The Manager manually adds the notes to the ATM cabinets
Balance Enquire	No requirements	Displays the current Balance, and information about last transaction

II. The construction of the ATM

The aim of this paper was to design a fully working replica of real ATMs that can be powered with 5v USB charger.

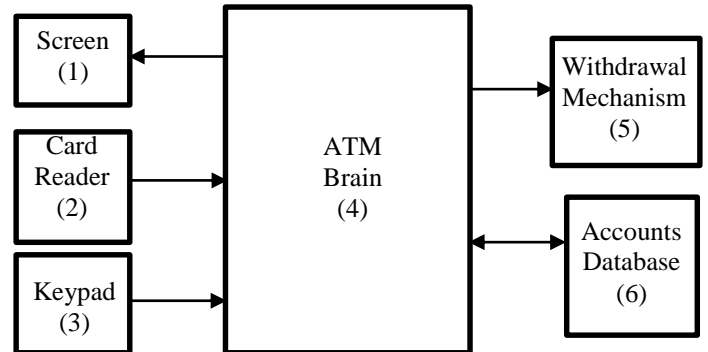


Fig. 2 General Block Diagram of the Automated Teller Machine (ATM)

The main electronic components used in the design are justified in the following paragraphs.

Starting with the screen (upper left block in Fig.2), there are many available options, including Thin Film Transistor (TFT), Organic light emitting diode (OLED) and Liquid Crystal Display (LCD), and since this ATM will only display text; hence the ASCII characters LCD is the best choice regarding the cost and complexity of programming. The 20x4 LCD module was used.

For the card reader block, there are generally two famous categories: magnetic tape and Radio Frequency Identification (RFID) technologies. The RFID was chosen, since its simpler to mount it to the ATM, also this technology does not require any physical contact between the reader and the card. The commonly known MFRC522 RFID module was used [7].

Moving to the keypad (the third block of fig. 2) which is the interfacing method between the ATM and the customer, there exists premade keypads that can be configured, however since two not adjacent sets of keypads are required (one for the numbers keypad, and the other for the screen options) therefore, using premade keypads would require more GPIO pins, hence two custom keypads were designed using 12mm push buttons and connected as one keypad.

The fourth and main block is the ATM brain which has the most variation in the choice of the used microcontroller (μc), the main dividing factors are: is compatible with the chosen components, can be easily programed, has enough storage space, has enough general purpose input/output (GPIO) pins, has a wireless network interface card (NIC). There are many existing μc s that satisfy the above requirements, however most of them are not available in the local stores, except for two famous modules namely the ESP32, and the ESP8266 manufactured by the Espressif company. These two modules are quite similar regarding the features of interest, however the dividing feature which has preferred the use of the ESP32 over the ESP8266, is in

the number of GPIO pins available for use. The ESP32 dev module was used [8].

The fifth block is designated for the withdrawal mechanism. This block represents most of the ATM's body (i.e., motors, sensors, gears and all other mechanical components) that work together to route the money notes out of the ATM (in the withdrawal process), and into the ATM (in the deposit process). The withdrawal mechanism used in this ATM is further discussed in the Hardware implementation section.

The sixth block (accounts database) represents the location where the customers' accounts are resident and how they will be accessed to fetch or update the relevant information. Many free online websites that offers APIs communicated databases, but each website has its drawbacks either in the amount of allowed memory or the displaying format. Another solution would require the creation of a local http website and host it on a computer, then the database of this website can be used to store the needed information about the customers, however for illustration purposes for the ATM services, this http website needs to be publicly accessed anywhere, hence a public domain named *atm.root.ly*, was registered from the Libyan Spider company, on which this http website was hosted. When this website is accessed the database of the customers is displayed in a table format as shown in Table 2

Table 2: The customer's table as displayed in The *atm.root.ly* domain, the table is only viewd for illustration purposes.

#	NAME	PIN	CASH	LAST CHANGE	LAST TIME
1	customer-1	2357	225	100	2020-12-13 13:15:48
2	customer-2	1966	480	-35	2020-12-13 11:27:48
3	customer-3	5681	555	--	--
4	customer-4	1416	335	60	2020-12-13 11:26:51

III. Hardware Implementation and design

Most of the ATM functions are not mechanical (ex: Balance Enquire, Change Pin, Transfer Money), and hence easy to implement. They only require lines of code to interface the user with the machine. However, two functions are purely mechanical (namely: Money withdraw, and Money deposit), and hence their implementation is not straight forward.

In real ATMs the withdraw function is implemented in a very complex way, with many of sensors and motors to route the money notes while checking their thickness for possibility of multi-note withdraw or fake note and handle the situation in each case (either return the excess notes to their cabinet or route the fake note to their cabinet).

Due to the complexity of the withdraw process used in real ATMs, and since this is a 5v operating ATM, a completely different decided idea is needed.

The proposed idea was to assume a cabinet with a stack of notes, and the withdrawal process will start from the top note using rotating rubber wheels. The first step is to design

an arm that will move in an arc using servo motor, at its wrist a small bidirectional dc motor is attached, this motor is further connected to two rubber wheels using bevel gears (to change the rotation direction by 90°). When withdrawing a single note, the top note will move outwards and the next top will move slightly due to friction, after the top note drops from the cabinet, the motor rotates in the reverse direction to restore the next top note to its starting position. The JX PDI-6209MG (servo motor) and the feetech fs90r (bidirectional motor) were used to build the wheeled arm. An illustration of this arm in the withdrawing process is shown in Fig.3.



Fig. 3 The wheeled arm while withdrawing

Now the mechanism for withdrawing from a single cabinet is finished, the same mechanism needs to be integrate to handle the case of multiple cabinets, an easy solution will be to design a wheeled arm for each cabinet but this is not an efficient use of resources, in addition to the excessive power requirements from these eight motors which cannot be satisfied from a regular 5v adapter. The another solution will require the use of mechanical encoding to transfer this wheeled arm to the intended cabinet, one way to implement this solution is by using the rack & pinion gears arrangement, the pinion is first attached to a stepper motor (28BYJ-48 stepper motor was used), this gear is further connected to the rack which was attached a slide drawer (linear equivalent of a ball bearing), the final implementation is outlined in Fig. 4, when the pinion rotates the wheeled arm moves to the left/right until it reach the intended money cabinet.

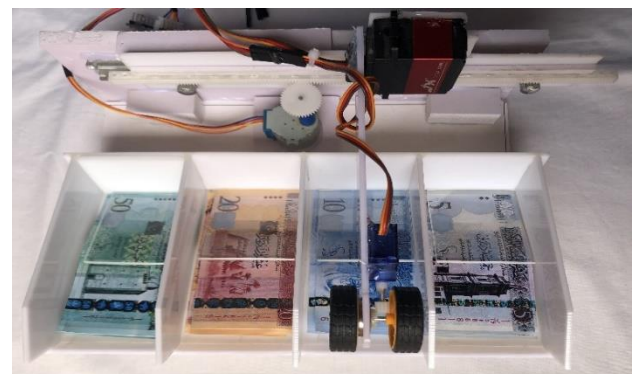


Fig. 4 Mechanical encoding when applied to the wheeled arm system of figure 3

When the intended number of notes is withdrawn, they are received on a small drawer (like the one used in assembly lines) then they are routed out of the ATM.

The deposit mechanism as implemented in real ATMs requires more sophisticated system, that first needs to accept the notes and then recognize them in addition to placing the notes into their cabinets. Due to the lack of needed sensors in the local stores, and the lack of additional GPIO pins that this mechanism requires, the mechanical implementation of this service was dispensed, however this ATM allows deposits added manually by the manager of the ATM.

When the withdrawing mechanism was solved, the next design step is to solve the shortage in the number of GPIO pins available for use in the ESP32 μ c. This μ c has a total of 23 different GPIO pins (Note: the used ESP32 μ c actually has a total of 31 GPIO pins, however 6 of them are used internally by the μ c for the integrated SPI chip, furthermore 2 more pins are used by the UART communication protocol, and all of these pins are not recommend for other uses as in [9], which leaves only 23 usable GPIO pins), while the components need a total of 37 pins, therefore encoding techniques must be implemented to solve this problem.

The first encoding technique (i.e. mechanical encoding) is already introduced in the hardware implementation section, this second encoding technique is the one applied to the keypad block. Since no two or more push buttons are to be pressed simultaneously, then each push button does not require a separate control pin to scene it. By the help of an encoder, the pressed button could be determined. For sixteen push buttons, a 16x4 encoder is required, such chip does not exist, however it could be implemented using two 8x3 encoder chips and three 2x1 Multiplexers, as shown in Fig. 5. This implementation saved a total of eleven pins by using only five pins (four pins were used to encode the 16

The last encoding technique was implemented between the LCD and the 28BYJ-48 stepper motor. The LCD uses six different pins (four data pins, register select (RS) pin, and enable pin). To communicate with the LCD the μ c sends the relevant instruction to the data and RS pins then it triggers the enable pin for the LCD to execute this instruction (i.e. the state of these pins is irrelevant to the LCD if the enable pin was not triggered). This property of the enable pin can be utilized to control both the LCD and the stepper motor in a technique named Address Latch Enable (ALE). This technique is famous among microprocessors with few GPIO pins, when reading or writing to memories [11]. In this case the LCD is analogous to the data lines of the memory, while the stepper motor is analogous to the address lines of the memory, this implementation saved three more pins by using only seven (six for the LCD, four of which are common plus one used to activate the latch), the ALE implementation is shown in Fig. 6. (Note: using stepper motor driver, and LCD with I^2C module can be helpful, however the drive requires a supply voltage of 8v or more to function whereas this is a 5v machine, using LCDs with integrated I^2C chip can dramatically saves 4 pins used to control the LCD, however such LCD is not available in the local store.)

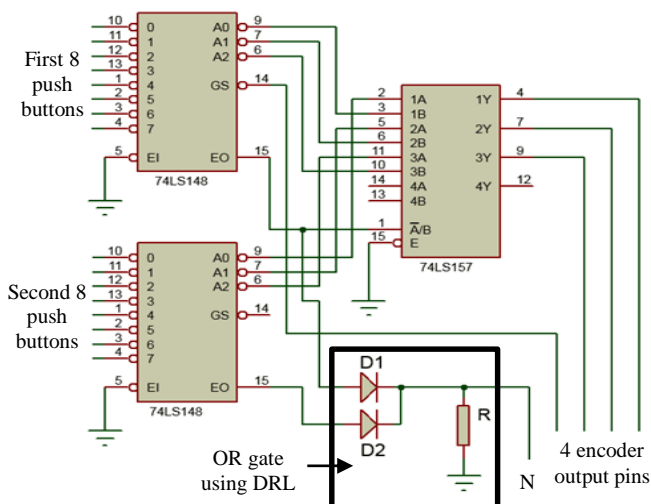


Fig. 5 logic circuit for encoding the 16 push buttons using four pins to encode the state and extra pin (N) to tell the null/active state.

push buttons, and extra pin for the null state) instead of total of sixteen [10].

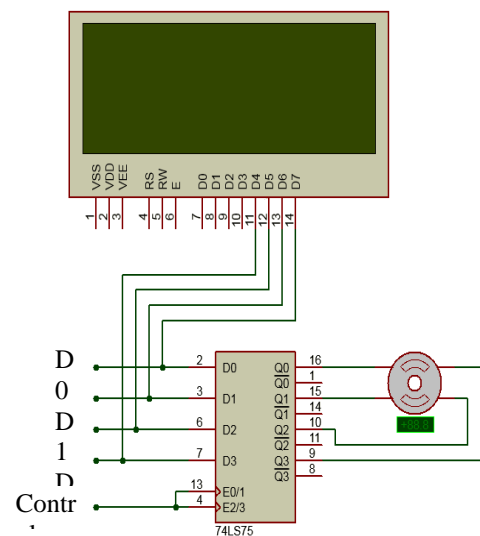


Fig. 6 logic circuit for the ALE technique used to share the bus between the stepper motor and 20x4 LCD screen.

IV. Programming And Software Implementation

The software part is divided into two sections, the offline section which regards the programming of the ESP32 microcontroller to control various ATM components while interfacing with the customer, and the online section which connects to the internet to fetch/update the information stored in the database of the virtual bank according to the customer's transaction.

Offline section: This section was written using the open-source Arduino IDE which resembles the programming with C & C++ [12]. This code can be divided into multiple blocks, where each block is dedicated for a specific function. The first block is a class that was written for the RFID card to make the read and write operations more natural and simpler, the main functions in this class includes the detection of the RFID card during the insertion and removal operations, in addition to verifying the identity of the customer, furthermore the class includes functions to communicate with the server via http messages to fetch/update the relevant information of the customer. The second block is allocated for the withdrawing option. When the customer is granted to withdraw the entering amount, its first divided among the available note using the famous withdrawing principle which states that: *at least one note to be dispensed from the lowest denomination available in the ATM*. The third block of code is dedicated to command the wheeled arm and the stepper motor to withdraw from the specified denominations, then the drawer is commanded to route the notes out of the ATM. The fourth block is specifically allocated for the LCD. It is written in FSM fashion, to display the correct screen at the correct time (hence the name Automated in the ATM), these screens includes options screens, comments to interact with the user, and screens that are triggered whenever an error happens (ex: accidentally removed card, wrong pin, insufficient balance.)

Online section: This section of code was written in the server side using the web-based programming languages PHP, MySQL, HTML, CSS. The front page of the domain (<https://atm.root.ly>) shows (in a tabular format) the information of the registered customers, this information includes the name of the customer, pin number, balance, and the last conducted transaction (Note: This information is included for illustration purposes, otherwise this information is sensitive and must be kept private), in addition to the customers' table, the page also shows the amount of each money denomination currently available in the ATM in a separate table. As the same time these tables are displayed the page contains PHP code that continuously monitors the URL of the page for an http message request sent by the ESP32, these messages share the same format of:

<https://atm.root.ly/?x=a&y=b&z=c>

This message is interpreted by PHP as:

There are three variables x , y , and z having the values a , b , and c respectively, by this way the ESP32 can tell the page to edit/get/add any data by specifying the right values to predefined variables. After each message sent by the ESP32, the page responds accordingly by sending either OK/NO message or the relevant information requested by the ESP32.

V. Communication and Security

When the ATM (Fig. 7) is first powered, it prompts a message to select the database domain, and the wireless

network (from a previously specified pool of networks). When the ATM successfully connects to the wireless network, the ATM starts by sending an http message requesting the available amounts of each money denomination resident in the ATM.



Fig. 7 Isometric image of the mini-ATM

Then the transaction loop of Fig.1, starts with the customer inserting his card. Then the RFID reader recognizes the card's UID number (this number is analogous to the MAC address used to identify the network interface cards (NICs)), which will be used to recognize the customer afterwards. The ESP32 then sends a http message to the web page requesting the customer's information associated with this UID, the web page responds by sending the relevant information (or a deny message if the customer is not registered) in a predefined format that can be understood by the ESP32. After this information is received the LCD displays a message asking the customer to enter his pin number, then the cycle continue as in the FSM of fig.1 between fetching and updating this information until the customer completes his transactions. [13]

One important service this ATM offers is ability to add new customer, if a new customer is to be added to the virtual database of this ATM, the Manager (special customer having this privilege) can (after inserts his card and entering his pin number) choose this option, the LCD then displays a message telling the Manager to remove his card and inserting the card of the new customer (any new RFID card compatible with the MFRC522 RFID reader module), the ESP32 then sends Http message to add this customer, the web page first checks if this card is not already registered (by checking the database of the customers), if so, a new account is added to the database with zero balance and a default pin of 0000 which can be changed later using the pin change service.

VI. Conclusion

The main goal of this project was to design and implement a mini ATM replica operating under 5v USB adapter. This prototype model will be of great interest (due to being noncommercial model of the actual ATM) specially to researchers for enhancing ATM services. Furthermore, it can serve as a first step to people with no prior experience in dealing with ATMs. Finally, this work is not innovative in the idea of designing a low voltage operating ATM, however it is innovative as being the first 5v model that resembles real ATM in a rational fashion.

There is a room to many enhancements to this ATM, the main enhancements with include: Adding a backup battery for both the ATM, and the network router in case of power outage, suction cubs to enhance the withdrawing mechanism, also the user's authentication can be enhanced by adding two factor authentication method and one-time password received on the customer's phone, or including biometrics indices.

References

- [1] Okokpujie, K., Noma-Osaghae, E., Okesola, O., Omoruyi, O., Okereke, C., John, S., & Okokpujie, I. P. (2018, June). "Integration of iris biometrics in automated teller machines for enhanced user authentication," In *International conference on information science and applications* (pp. 219-228). Springer, Singapore.
- [2] Wang, Y., Zhang, Y., Sheu, P. C., Li, X., & Guo, H. (2010). "The formal design model of an automatic teller machine (ATM)," *International Journal of Software Science and Computational Intelligence (IJSSCI)*.
- [3] Khalifa, S. S., & Saadan, K. (2013). "The formal design model of an automatic teller machine (ATM)".
- [4] B. Saranraj, N. S. P. Dharshini, R. Suvetha, and K. U. Bharathi, "ATM Security System Using Arduino," presented at the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS).
- [5] Brijesh Singh, 2 Factor Authentication for ATM Machine Project, 25 October 2019, accessed on May, 2020, <https://create.arduino.cc/projecthub/skyline/2-factor-authentication-for-atm-machine-project-969c27>
- [6] Sooraj P, Bony Mons, Dr. Jinsa Kuruvilla, June (2019), "IoT Based Vending Machine with Cashless Payment," *International Research Journal of Engineering and Technology (IRJET)*, vol. 6, no. 6.
- [7] What is RFID? (no date), accessed on May 2020, available at: <https://lastminuteengineers.com/how-rfid-works-rc522-arduino-tutorial/>
- [8] ESP32. (no date), accessed on May 2020, available at: <https://www.espressif.com/en/products/socs/esp32>
- [9] ESP32 Pinout Reference: Which GPIO pins should you use? (no data), accessed on June 2021, available at: <https://randomnerdtutorials.com/esp32-pinout-reference-gpios/>
- [10] Mano, M. Morris and Ciletti, Michael D. *Digital Design*. Upper Saddle River, NJ: Pearson Prentice Hall, cop. 2013
- [11] Brey, Barry B. *The Intel microprocessors: 8086/8088, 80186/80188, 80286, 80386, 80486, Pentium, Pentium Pro processor, Pentium II, Pentium III, Pentium 4, and Core2 with 64-bit extensions: architecture, programming, and interfacing*. Upper Saddle River, NJ: Pearson Prentice Hall, cop. 2009
- [12] Arduino ide. (no date), accessed on May 2020, available at: <https://www.arduino.cc/en/software>.
- [13] Gyamfi, N. K., Mohammed, M. A., Nuamah-Gyambra, K., Katsriku, F., & Abdulah, J. D. (2016). "Enhancing the security features of automated teller machines (ATMs): A Ghanaian perspective," *International Journal of Applied Science and Technology*.

Implementation of the Computer Numerical Control Milling Machine

¹ Nabil T. Alelwani, ² Abdulhamed M. Hwas

^{1,2} Department of Mechanical and Industrial Engineering, University of Tripoli, Tripoli, Libya

¹n.olwani@uot.edu.ly, ²A.Hwas@uot.edu.ly

ABSTRACT

The purpose of this project is to implement a fully functional 3-axis Computer Numerical Control (CNC) machine that is capable of milling a work piece with the sequence of steps carried out to implement the basic principles of a CNC machine. The objective is to build a Computer Numerical control milling machine using Open Source software and a microcontroller system. Here, a prototype of a CNC milling machine is made to engrave and cut shapes and data on a rectangular coordinate system. The machine contains 4 stepper motors that control the 3 axis X, Y and Z (Y-axis requiring 2 steppers) held in place with a combination of a wooden frame and 3 Dimension printed parts that were designed, a printed circuit board was designed and used with a microcontroller to move the stepper motors after interpreting the g-code that was streamed from a USB and the cutting tool as a variable speed spindle. The prototype CNC worked well and some designs and models were made successfully.

Keywords:

Computer numerical control, milling machine, electronic design automation.

I. Introduction

The motivation for this project was derived from a life-long passion for combining the three major fields (Mechanical, Electrical and Computer Engineering) together to implement complex systems in which the three aspects are complementing one another. The best demonstration was to build by hand one of these machines that helped us take a leap forward in the field of automation. CNC machining is a manufacturing process in which pre-programmed computer software dictates the movement of factory tools and machinery. The process can be used to control a range of complex machinery, from grinders and lathes to mills and routers. With CNC machining, three-dimensional (or more) cutting tasks can be accomplished in a single set of prompts. In the modern-day, the need to produce large quantity products with high quality and low production cost is higher than ever, this can be achieved by the machines which are controlled by computers to make work easier with ultra-precision and less human caused errors. These are commonly known as CNC Machines. Computer Numerical Control (CNC) machining is a process used in the manufacturing sector that involves the use of computers to control machine tools. Under CNC machining, machine tools function through numerical control. A computer program is customized for an object and the machines are

programmed with a language called G-code that essentially controls all features like feed rate, coordination, axis motions, location and speed.

The next section of this paper will go over the subsystems that drive the CNC. Then cover the work on the implementation and the steps required to achieve the objective as well as show the results of some work pieces; and finally the conclusion.

II. Computer Numerical Control Machine

Computer Numerical Control Machine has revolutionized the manufacturing industry. Whether it is a small manufacturing unit or a large global company, CNC machines find applications in almost all types of industries. That's because when it comes to producing complex parts in metal or any other material, these computers controlled machine tools are ideal because of their high levels of accuracy, precision and speed. So let's take a look at the various industries where CNC is used. Metal Removal Applications: CNC machines are extensively used in industries where metal removal is required. The machines remove excess metal from raw materials to create complex parts. A good example of this would be the automotive industries where gears, shafts and

other complex parts are carved from the raw material. CNC machines are also used in the manufacturing industries for producing rectangular, square, rounded and even threaded jobs. Metal Fabrication Industry: Many industries require thin plates for different purposes. These industries use CNC machines for a number of machining operations such as plasma or flame cutting, laser cutting, shearing, forming and welding to create these plates. Electrical Discharge Machines: (EDMs) remove metal from the raw material by producing sparks that burn away the excess metal. EDM machining through CNC automation is carried out in two different ways; first through Wire EDM and second through Vertical EDM [7]. Besides these industries, CNC machines are also used in the woodworking industries for various operations like drilling and routing.

III. Computer Numerical Control Implementation

CNC can be divided into three subsystems, Firstly, the mechanical subsystem gets necessary control signals from the electronics sub-system, which ultimately results in desired actuation of motors. Secondly, the electronics subsystem gets command or a set of a commands from the software subsystem and generates controls for the mechanical subsystem. Thirdly, the software subsystem which is an Arduino mega in this project which interprets the g-code instructions into the required movement, these instructions are usually generated from the CAD/CAM software design and translates into commands.

A. Mechanical Subsystem

The CNC will be built around the Cartesian coordinate system, in which every axis (x, y and z) are all perpendicular to each other. The frame will consist of a wooden base alongside a bunch of 3D-printed parts that are responsible for housing the stepper motors, and the smooth rods the whole mechanism will slide along as well as the lead screw that will generate the required torque from the stepper motors.

- Designing and modeling the parts:

Autodesk Fusion 360 was used to model the required parts, "Fusion 360 helps students and educators prepare for the future of design. It's the first 3D CAD, CAM, and CAE tool of its kind, connecting your entire product development process into one cloud-based platform" [13]. With this software, could be a draw, model and render objects that can be converted into files to use on a 3D printer. It utilizes cloud technology in a way that not many other CAD programs do. Instead of using copious amounts of processing power on your computer, Fusion 360 uses the Autodesk servers so you don't necessarily need an extremely powerful computer if you are carrying out complex tasks. By acting as a 'resource multiplier', the cloud technology is able to utilize much more power than if you were simply using your own machine. You can also access all your files and data via the cloud too. This makes it better for accessing files on the move and you don't need to physically take them with you on a storage device.

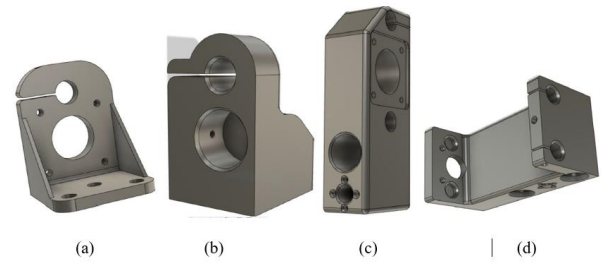


Figure 1. Modeling the parts, a) Y Stepper motor housing, b) The rest of the lead screw and the smooth rod in the Y-axis, c) The rest of the lead screw and the smooth rod in the X-axis, d) The spindle carriage with the z motor housing on top

Figure 2. housing on top

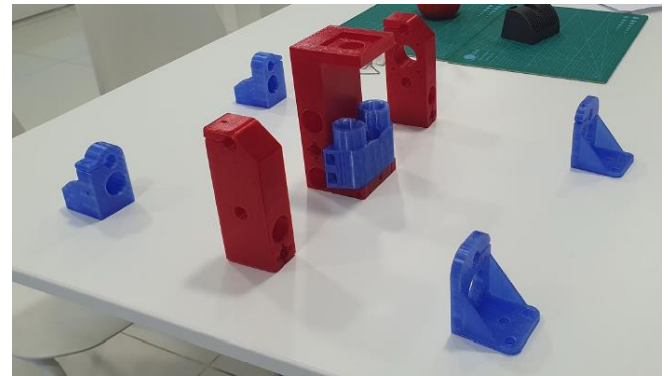


Figure 3. The designs after being printed

- Printing material:

The filament used for printing these models is PETG. PETG filament is a Glycol Modified version of Polyethylene Terephthalate (PET) and it's an extra tough 3D print material. This is an extremely high strength filament and can achieve very sturdy and strong prints. It has very low shrinkage, making this perfect for larger flat surfaces. PETG is a perfect alternative to ABS and PLA, offering higher strength, lower shrinkage, and a smoother finish [14]. PETG is the perfect filament to combine strength and ductility, which is why it's used in so many mechanical parts and robotics. It has great chemical resistance with good water, acidic and alkalic resistance.

Mechanical property:

- Elongation at Break: 21.5%
- Tensile Strength: 50 MPa
- Bending Strength: 68 Mpa
- Impact Strength IZOD: 8 KJ/m2

- The guide rail system:

This system will consist of **A)** two smooth rods on the y-axis 500mm long that are responsible for carrying the whole mechanism which are moved by the lead screw that are being turned by the two stepper motors, **B)** two more smooth rods for the x-axis 300mm long that are being moved by the

y-axis and is responsible for carrying the spindle and the z motor, C) two last smooth rods for the z-axis that will move the carriage holding the spindle.

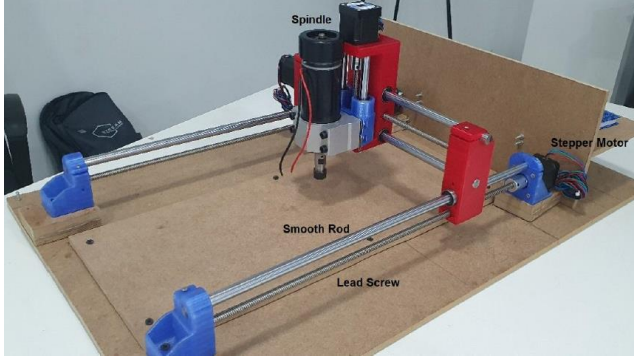


Figure 4. CNC machine frame prototype constructed for this project

B. Electrical and Electronic subsystem

This is the subsystem responsible for generating control signals to the stepper motors which guides the motion of the tool path in each direction or axis and it consists of the main following components:

- Power Supply Unit (PSU):

This is the unit that powers the whole system (apart from the Arduino), it basically converts AC voltage (220~240V) to a DC 24V for the motors.

- Stepper Motors:

As it is known, the stepper motor is a brushless, synchronous electric motor that converts digital pulses into mechanical shaft rotation. The motor's position can then be controlled to move and hold at one of these steps without any feedback sensor (an open-loop controller). When working with stepper motors a few characteristics need to be known for it to be controlled accurately, the most important is how many steps does the motor need to perform in order for it to move 1mm or (Step/mm). In this project, a 1.8° step angle motor was used that requires it to do 200 steps to complete one revolution (1. 8° × 200 = 360°), and the lead screws that were used have 2mm/revolution pitch, that gives us 200 Step per 2mm or 100 steps/mm at full step, but usually, stepper motors are operated with micro-stepping which results in more accuracy A 1/8 micro-stepping was chosen for this which increases our step/mm by 8 giving us 800 steps/mm at 1/8 micro-stepping.

The motors used in this project have a step angle of 1.8 degrees. The rotation of these motors is controlled by the electric current which pours into X coil, X' coil, Y coil and Y' coil using full-step mode as shown in Table I.

TABLE I. FULL STEP SEQUENCE OF STEPPER MOTOR

Step	X	X'	Y	Y'	Angle
1	0	1	0	1	0.8°
2	1	0	0	1	1.8°
3	1	0	1	0	3.6°
4	0	1	1	0	5.4°
5	0	1	0	1	7.2°

- Stepper Motor Specifications:

The main key specification we wanted was the torque required by the motors in the configuration we have (see figure 4.8), using Oriental motor [15] online calculator tool which matched the manual calculation according to Table II, and the following equation:

$$T_L = \left(\frac{FP_B}{2\pi\eta} + \frac{P_B F_O \mu_0}{2\pi} \right)$$

With a safety factor of 2.

TABLE II. MOTOR RELATED INFO

Total mass	5 kg
Force $F = mg$ (F)	$5 \times 9.81 \approx 50$ N
Initial force $F_0 = F/3$	~ 16.5 N
Friction coefficient of lead screw μ_0	(Steel) 0.8
Lead Screw Diameter	8 mm
Total length	500 mm
Lead (pitch) PB	2 mm /rev
Lead Screw Efficiency η	$\sim 30\%$
Operating speed (Fixed speed)	20 mm / s
Acceleration/Deceleration	5 s
Stopping accuracy	0.1 mm

Which yields a loading torque of ~ 0.0457 N·m, and an acceleration torque of ~ 0.0915 N·m. The stepper motor that was chosen has a loading torque of 0.59 N·m.

- Stepper Motor Driver IC (DRV8825) as shown in Figure 1 [11]. This is the IC that drives the motor with the 24V, it receives from the PSU (and it also powers itself with a built-in voltage regulator).
- The Spindle: The tool head that is responsible for cutting /milling (machining)
- Limit Switches: These switches are placed at the edges of the working space to give the CNC a sense of awareness of the edges, sort of a closed-loop system.
- Arduino: This is the brain of the CNC that interprets the G-Code into actual movement.

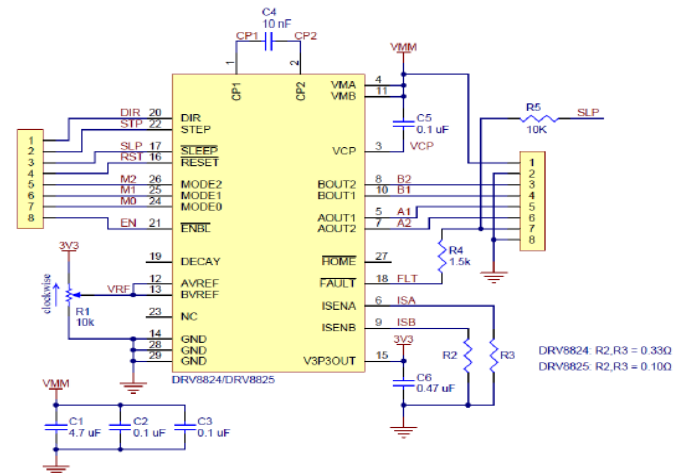


Figure 5. DRV8825 circuit diagram.

- Printed Circuit Board (PCB):

A printed circuit board (PCB) mechanically supports and electrically connects electronic components or electrical components using conductive tracks, pads and other features etched from one or more sheet layers of copper laminated onto and/or between sheet layers of a non-conductive substrate. Components are generally soldered onto the PCB to both electrically connect and mechanically fasten them to it.

- Electronic Design Automation (EDA):

Electronic design automation, also referred to as electronic computer-aided design (ECAD), is a category of software tools for designing electronic systems such as integrated circuits and printed circuit boards. The tools work together in a design flow that chip designers use to design and analyze entire semiconductor chips [16]. KiCad was used in this project. KiCad is a free software suite for electronic design automation. It facilitates the design of schematics for electronic circuits and their conversion to PCB designs. It features an integrated environment for schematic capture and PCB layout design [17].

The Process of designing a PCB starts with imagining how all the components go together to the very basic ones like switches, capacitors, connectors etc..., Then we start placing these components in the Eeschema interface (Note that some components aren't available by default (like the DRV8825 IC) which require us to actually design those components from scratch following their datasheet carefully).

After finishing the schematic and checking all the wirings are correct, the next step is to assign a footprint (which is the physical representation of the component) for every component. The next step is to use the "PCBNew" tool to arrange the components and route the copper trace like our schematic, with clutter as shown in Figure 5. Where every component's footprint is tangled to every other component's footprint according to the schematics. And after working my way around placing the components and routing the copper trace between them as shown in Figure 6.

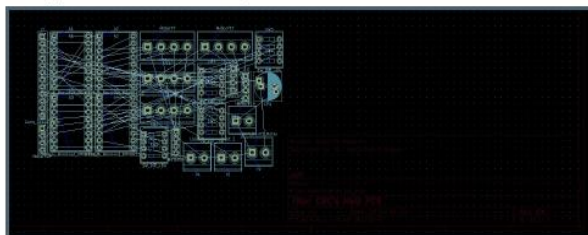


Figure 6. KiCad layout at the beginning of designing the PCB footprint

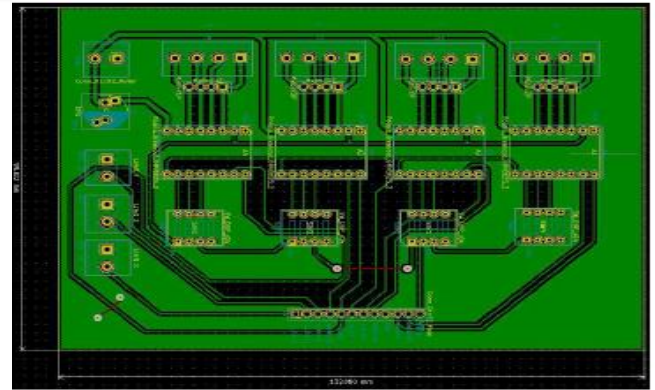


Figure 7. Final PCB layout after routing all the traces and placing the components in their desired position

Where the green parts are copper filled traces, note that wider traces were used for carrying power and thinner ones were used for logic signals. Power trace width: 48mils, motor trace width: 36mils, logic trace width: 18mils*. *The imperial system was used here because it's considered the standard in PCB design where (1mil = 0.001 inch). A 3D render of the design (see Figure 7).

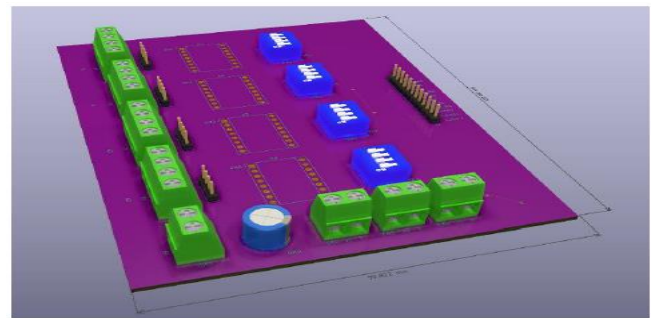
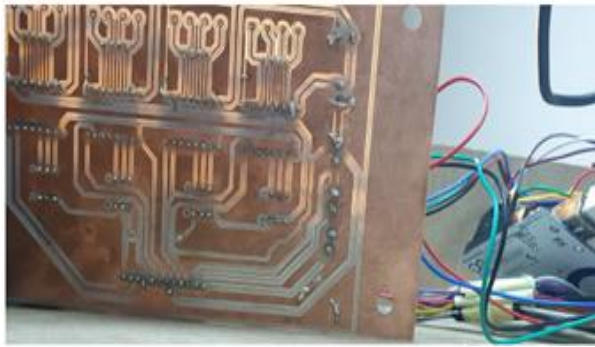


Figure 8. A 3D render of how the PCB should look like

The design was then be exported into Gerber files to be CNCed "Milling". PCB after a lot of failures to get it right, a CNC was used to mill this PCB and drill the holes. Where every component's footprint is tangled to every other component's footprint according to the schematics. And after working my way around placing the components and routing the copper trace between them (see Figure 8).



(a)



(b)

Figure 9. PCB after soldering and wiring the components, a) components on the PCB, b) A bottom view of the PCB

C. Software subsystem

This is the system that takes the g-code generated from a CAD/CAM software and interprets it as commands sent to the electronic subsystem that acts accordingly.

- The Firmware:

GRBL firmware was used in this project [18]. Grbl is a high-performance software for controlling the motion of machines that move, that make things move, and will run on a straight Arduino. If the maker movement was an industry, Grbl would be the industry standard. Most open-source 3D printers have Grbl in their hearts. It has been adapted for use in hundreds of projects including laser cutters, automatic hand writers, hole drillers, graffiti painters and oddball drawing machines. Due to its performance, simplicity and frugal hardware requirements Grbl has grown into a little open-source phenomenon.

- The Software:

This is the part that generates the G-code and streams it to the Arduino, OpenBuilds [19] was used in this project. Open Builds control is an application for connecting to, and controlling a CNC, Laser or Plasma. Its features are:

- Interface with, and Jog your machine
- Run GCODE Jobs
- Set Zero coordinates
- Integrate with cam.openbuilds.com
- Flatten/Surface your spoil board/stock
- Help with Firmware configuration

OpenBuilds CAM is a web-based application for converting SVG, DXF and Bitmap drawings to GCODE for use with your CNC, Laser or Plasma.

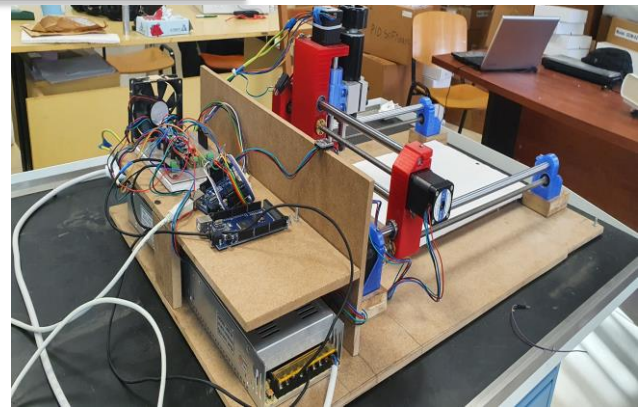


Figure 10. Computer Numerical Control Milling Machine prototype constructed for this project.

IV. Results and discussion

Figure 9 shows the CNC Milling Machine prototype constructed for this project. And Figure 10 shows the desired object. The design was made using OpenBuilds CAM by converting the text into G-code with these settings:

- Type of cut: CNC Vector (No offset)
- Drill bit diameter: 1.0mm
- Z safe height: 15.0mm
- Cut depth per pass: 2.0mm
- Cut depth final: 2.0mm
- X/Y speed: 1200mm/min
- Z speed: 600mm/min

TABLE III. CNC WORKING AREA

Axis	Safe Travel Distance (mm)
X axis travel	185
Y axis travel	350
Z axis travel	55

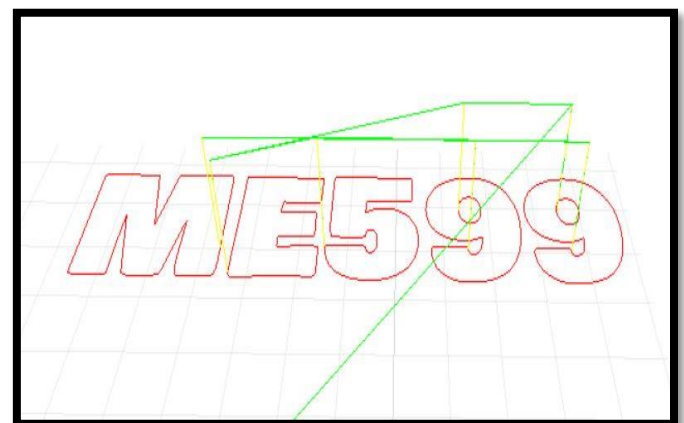


Figure 11. An example object created with OpenBuilds CAM.

And these are the results:



Figure 12. The output of the CNC with different fonts.

The results show that the CNC is functional with a low cost budget, but with a little inaccuracy in the X axis which is due to the faulty linear bearings that causes some jittering and slipping.

V. Conclusion

Implementation of the CNC started by having a rough sketch of how it should look like and designing the frame and the guide rail system in a Cartesian fashion to ensure that the three axes can move freely and steadily. A 3d printer was used to print the required parts and then assembled them to their final shape.

Then in the electronic part and by understanding a stepper motor works and the circuit of its driver IC all the proper components were picked and PCB was designed to house these all components in a well-ordered way that integrates with the microcontroller. PCB design was then exported into Gerber files and using a CNC. After many failed attempts to mill it to get all the traces correctly with no short or open circuits (which had a lot of errors due to CNC inaccuracy) and after drilling the pins, the components were soldered and wired according to the schematic.

Finally, for the firmware, GRBL was used in the Arduino to interpret incoming g-code. And OpenBuilds was used as the software that connects to it to stream. The g-code, that was generated using OpenBuilds online CAM tool. As future work, it suggests using linear rails instead of the smooth rods as they can live for longer and can withstand harsh conditions and using a belt-driven mechanism instead of a leadscrew for the driving system for more and efficiency, also a DC motor with an encoder would be better than stepper motors to achieve a closed-loop system.

References

- [1] En.wikipedia.org. 2020. *Metal Lathe*. [online] Available at: <https://en.wikipedia.org/wiki/Metal_lathe> [Accessed 31 December 2020].
- [2] Prototech Asia. 2020. What Is CNC Machining Technology ? History And Definition. [online] Available at: <<https://prototechasia.com/en/plastic-cnc-machining/history-cnc/>> [Accessed 31 December 2020].

- [3] Make: DIY Projects and Ideas for Makers. 2020. Punch Tape Programmable Metal Mill From 1952 | Make:. [online] Available at: <<https://makezine.com/2012/04/02/punch-tape-programmable-metal-mill-from-1952/>> [Accessed 31 December 2020].
- [4] N. Alelwani, A. Hwas, M. Makhlof and A. Alwahedi "Implementation of the Basic Principles of a Computer Numerical Control Machine" in *International Conference on Technical Sciences, ICST 2019*, Tripoli, Libya, March, 04-12, 2019
- [5] Prusa Printers. 2020. *Prusa I3 3D Printer - Prusa Printers*. [online] Available at: <<https://blog.prusaprinters.org/prusa-i3/>> [Accessed 31 December 2020].
- [6] Michael Mattson, "CNC Programming: Principles and Applications", Book 2nd edition, 2010.
- [7] B. Jayachandriah, O.Vamsi Krishna, P.Abdullah Khan and R. Ananda Reddy, "Fabrication of Low Cost 3-Axis Cnc Router", *International Journal of Engineering Science Invention*, volume 3 Issue 61, June 2014.
- [8] B.S. Pabla and M.Adithan, "CNC Machines", Book, 1994.
- [9] Notesandsketches.co.uk. 2020. *Motion Control Using Screws*. [online] Available at: <http://www.notesandsketches.co.uk/Motion_control.html> [Accessed 31 December 2020].
- [10] S.K. Pillai, "A First Course on Electrical Drives", Book, 2004.
- [11] Pololu.com. 2020. *Pololu - Schematic Diagram For The DRV8824/DRV8825 Stepper Motor Driver Carrier..* [online] Available at: <<https://www.pololu.com/picture/view/0J4222>> [Accessed 31 December 2020].
- [12] Sundar Pandian and S. Raj Pandian, "A Low-Cost Build-Your-Own Three Axis Cnc Mill Prototype", 2014.
- [13] [13] Design Academy. 2020. *Getting Started With Fusion 360*. [online] Available at: <<https://academy.autodesk.com/getting-started-fusion-360>> [Accessed 31 December 2020].
- [14] Airwolf 3D. 2020. *3D Printing With PETG: Tips And Tricks - Airwolf 3D Printers*. [online] Available at: <<https://airwolf3d.com/2015/12/05/3d-printing-with-petg-tips-and-tricks/>> [Accessed 31 December 2020].
- [15] Oriental Motor U.S.A. Corp. 2020. *Ball / Lead Screw Sizing Tool*. [online] Available at: <<https://www.orientalmotor.com/technology/motor-sizing-calculations.html>> [Accessed 31 December 2020].
- [16] En.wikipedia.org. 2020. *Electronic Design Automation*. [online] Available at: <https://en.wikipedia.org/wiki/Electronic_design_automation> [Accessed 31 December 2020].
- [17] Kicad-pcb.org. 2020. *Kicad EDA*. [online] Available at: <<https://www.kicad-pcb.org/>> [Accessed 31 December 2020].
- [18] GitHub. 2020. *Gnea/Grbl*. [online] Available at: <<https://github.com/gnea/grbl/wiki>> [Accessed 31 December 2020].
- [19] Software.openbuilds.com. 2020. *Openbuilds Software - FREE Software For CNC Control: Openbuilds CONTROL And Openbuilds CAM*. [online] Available at: <<https://software.openbuilds.com/>> [Accessed 31 December 2020].

Modeling of Pneumatic Air Muscles for Rehabilitation Robotic Systems

¹ Ayah Hasan Alwahedi, ² Azeddien Kinsheel

^{1,2} Mechanical Engineering Department, University of Tripoli, Tripoli, Libya

¹ AyahAl1996@gmail.com, ² a.kinsheel@uot.edu.ly

ABSTRACT

Keywords:

Air Muscles, Fuzzy Logic, Least Square, Robotics, Rehabilitation.

Pneumatic Artificial Muscles are one of the famous linear actuators in bio-inspired robotics. They can generate a relatively high linear force considering their form factors and weights. Furthermore, PAMs are inexpensive when compared with traditional electromagnetic actuators and also inherently light and compliant. The movement of Pneumatic Artificial Muscle is soft resembling the human muscle. This feature makes the use of the actuator for rehabilitation devices and prostheses very appealing. In this paper, a model of Pneumatic Artificial Muscle is fabricated and tested. The dynamic model of the fabricated system is investigated. Two methods were used: Least Square method and Fuzzy Logic method. Our goal is to achieve a model with minimum error and force difference. The experimental test results are presented and discussed.

I. Introduction

Nature has always been a source of inspiration for engineering designs. As it is desired to develop an actuator that allows for the mimicking of the Human movement, the choice of the human muscle as the basis of design is the most reasonable. This human-like actuator is called a Pneumatic Artificial Muscle (PAM). The air muscle itself was originally developed by McKibbens to aid his paralyzed daughter in regaining some movement in her hand, [1].

Due to their lightweight, PAMs are: suitable for applications with limited spacing; compliant, easing control; safe, making its use possible in human interaction applications as a low-cost efficient actuator. To benefit from the outstanding advantages of the pneumatic actuators such as high softness and low weight-to-force ratio, efficient control of the actuator force as well as its displacement is essential. Attaining a comprehensive model with a satisfactory accuracy in the entire course of the muscle is the most important challenge regarding utilization of the pneumatic artificial muscle in a wide range of the applications, [2]. PAM is powered with pressurized air which makes them environmentally friendly. The possible implementations for the Air muscles are countless, as they can safely work hand in hand with humans. Such as, Pneumatic 3-DOF robot, PAM-actuated robot arm system, 2-DOF planar robotic manipulator, PAMs used in modal parameters of bridges, Parallel-kinematic hexapod

tool, Hybrid robot for safe human-robot interaction, Industrial gripper, and a Positioning system, [3].

A number of mechanisms have been designed and implemented with the use of PAMs even though a reliable model is yet to be developed. Despite the lack of a model, improvements in the field are being made every year. All the benefits provided by the McKibbens Air muscles make it understandable why this field developed a lot of interest over the years. Several scientists followed in McKibbens' footsteps. A few scientific breakthroughs made by these scientists such as clinical trials were published for a muscle made of pure rubber latex covered in a double helical weave and operated similarly to a human muscle, in the 1960s. Rubbertuators developed by Bridgestone Company in the 1980s & air muscles developed by Shadow Robotic Company in the 1990s [4]. The latter being the most commonly used and is associated with almost all humanoid applications that have been developed recently. As in all fields, there are competitive companies with similar designs such as Merlin Humaniform and a few others, [4].

The first attempt at developing a theoretical model was called the Basic formula established by a scientist named Gaylord in 1958, [4]. This model was the base for all the models developed in the years to follow. The interest in the modeling of PAMs died out for around 30 years until the 1990s where people became more environmentally aware and new sources of energy were desirable. In 1996, Chou and Hannaford

developed a model that considered the effect of bladder and braid thickness, [4]. They also gave weight to the effect of friction on the force calculation, and considered the muscle as a conic shape instead of a perfect cylinder. In 2000, a model by Tondou and Lopez explored the effect of the conic shape and the friction effect, at the same time, a model by Tsagarakis and Caldwell also considered the effect of assuming the muscle had a conic shape, [4]. This is differed from Tondou and Lopez work, in that they not only explored the effect of the shape of the muscle but also the bladder elasticity. Klute and Hannaford focused on the effect of bladder elasticity also. in 2003 Davis, Tsagarakis, Canderle, and Caldwell found that the fiber extension depends on the pressure and the muscle length. All the previous models had assumed that the braid length was constant; to overcome this issue an expression for variable braid length was established. Further improvement to the model is still possible, although a breakthrough in this field hasn't been found since 2003.

In this paper, a PAM to be used for rehabilitation robotic system is presented. The dynamic model of the fabricated test PAM is experimentally investigated. The aim of this work is to obtain a simplified model with an accuracy that is sufficient for the intended application. Two different approaches to obtain the dynamic model are considered i.e.; Least Square method and Fuzzy logic method. In the following sections the description of the system, its theoretical model and the experimental modeling are presented.

II. Principle and Modeling of PAMs

PAMs are described as expandable chambers that are powered by pressurized air. The chamber, also called the bladder, is surrounded by a braided sheath. The braided sheath allows the bladder when pressurized to expand equally along the muscle length in the radial direction. This causes the actuator to contract in the axial direction as shown in Figure (1). The contraction of the muscle is proportionate to the amount the diameter is forced to increase. The air muscle must be initially stretched in order for a contraction to occur when it is activated. Mechanical power produced by the contraction force is transferred to the load via the end fittings.

Figure 13. contracted state with an External Load and an External Force.

The PAM is theoretically believed to obey three simple Laws, [4].

Law 1: *An increase in Pressure brings about an increase in the Contraction Ratio.*

Law 2: *An increase in Load applied results in a smaller Contraction Ratio.*

Law 3: *An increase in Pressure increases the resulting Force*

Figure 1. It should be mentioned that the maximum contraction occurs when there is no load applied and the volume is maximum.

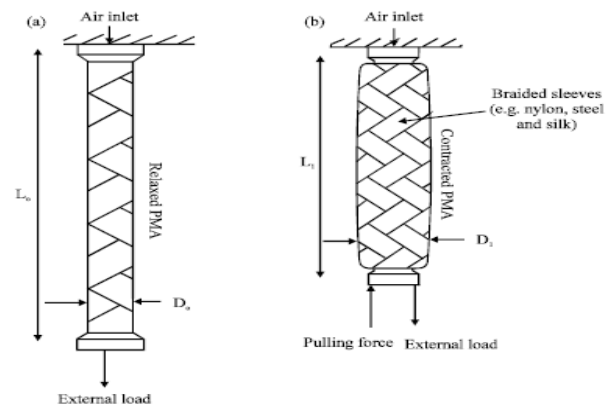


Figure 1. A PAM in relaxed state with an External Load and A PAM in

When paired PAMs have the ability to mimic the motion of a human limb, known as antagonistic setup, Figure (2). This is basically two muscles that contradict each other. As one muscle expands the other contracts. This set up is extremely beneficial in robotics and other human like applications where movement of a joint is required.

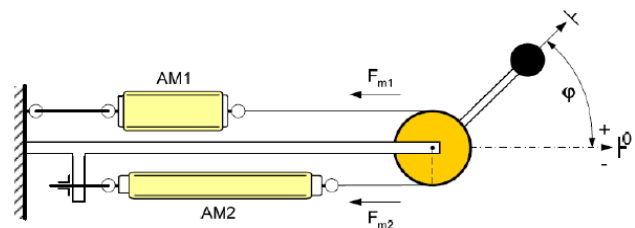


Figure 2. Antagonistic Set up of Pneumatic Air Muscles showing the movement of a Pulley.

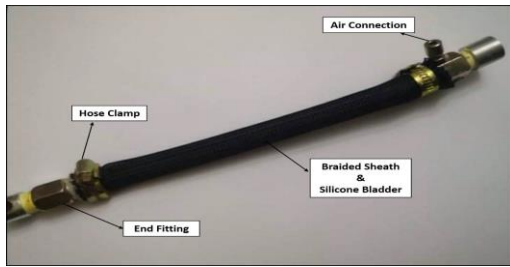
The required components to construct a pneumatic artificial muscle are simple and inexpensive, [5]. A variety of options are available when selecting the materials and lengths used for the bladder and braided sheath. Below is a list of components required to construct the PAMs.

- Silicone bladder
- Braided sheath
- End fittings
- Hose clamps
- Air connection
- Air compressor

When assembled these simple components are used to construct and operate the PAMs. Figure (3) shows a constructed PAM with air inlets. In order to construct a PAM, these instructions are followed, [5]: Cut out 10-inch-long piece from the mesh and silicone tubes. Thw silicone bladder is inserted into the mesh outer case, this must be done delicately. The end fitting is inserted into the bladder and a hose clamp is used to secure it in place. An air connection is

placed allowing the pressurized air to be supplied to the chamber.

Figure 3. Constructed Air Muscle with air inlet valve.



A. Theoretical Models

Over the years, several theoretical models were developed and examined for PAMs. Among these models is the improved model proposed by Tsagarakis and Caldwell [6] which takes the conic shape of the actuator and the force of linear elasticity into account, i.e;

Equation 1

$$F = p \frac{b\pi D_o^2 \cos^2 \theta \sin \theta + \frac{2}{3} b_1 \pi D_o^2 (\cos^3 \theta - \cos^3 \theta_{min})}{2bsin\theta + 4b_1(\cos\theta - \cos\theta_{min}) - p \frac{\pi D_o^2 \sin^2 \theta}{4} + K_{Rubber}(b \cos \theta - L_{Rubber})} + \pi p \left[D_o t_k \left(2 \sin \theta - \frac{1}{\sin \theta} \right) - t_k^2 \right]$$

When:

$$\theta > \sin^{-1} \frac{D_{cap}}{D_o}$$

And equals:

Equation 2

$$F = p \frac{b\pi D_o^2 \cos^2 \theta \sin \theta + \frac{2}{3} b_1 \pi D_o^2 (\cos^3 \theta - \cos^3 \theta_{min})}{2bsin\theta + 4b_1(\cos\theta - \cos\theta_{min}) - p \frac{\pi D_{cap}^2}{4} + K_{Rubber}(b \cos \theta - L_{Rubber})} + \pi p \left[D_o t_k \left(2 \sin \theta - \frac{1}{\sin \theta} \right) - t_k^2 \right]$$

When:

$$\theta < \sin^{-1} \frac{D_{cap}}{D_o}$$

Where:

θ	Braid angle
P	Absolute pressure
b	Braid strand
D_{cap}	Diameter of the End Cap
D_o	Theoretical diameter of the muscle when $\theta = 90^\circ$
θ_{min}	Minimum braid angle theoretically assumed to be 20°
t_k	Thickness of the bladder and braided sheath
K_{rub}	Linearized radial actuator elasticity
L_{rub}	Length of Rubber

The first two terms in Equation (1) represent the force equation derived by Tsagarakis and Caldwell, based on the force profile on the conic shape. The third term, also derived by Tsagarakis and Caldwell focuses on the effect of the bladder elasticity on the force developed. To further improve the model, Tsagarakis and Caldwell utilized the discoveries

made by Chou & Hannaford to consider the effect of the braid thickness on the force developed. This finding was integrated into Tsagarakis and Caldwell Equation, the final term in equation 1.

Tsagarakis and Caldwell compared this theoretical model with the experimental data obtained from extensive tests. According to their results the calculated force of the new model is closer to reality than the model proposed by Chou and Hannaford, where the actuator was considered a cylindrical shape and not a conic shape. They claimed that the model is particularly accurate in the high and mid contraction areas. And it is in general 30%-50% more accurate than other models. The error remains significant especially in low pressure areas - about 10 % error is observed.

There are several factors that contribute to the significance of the error, among them is the uncertainty in parameters and the nature of pneumatic systems and their sensitivity to temperature and humidity, also the flexibility of the tubes and the variation of the braid angle as the muscle shape changes. Furthermore, the theoretical model is highly nonlinear in its physical parameters which make it difficult to identify those parameters using classical identification methods. Therefore, and in order to obtain a simplified and yet accurate model, an experimental study of the muscle behavior is considered.

III. Experimentation

An Experimental model is based on acquired data from sensors, such as position, force and pressure sensors. This section discusses the means used to find the experimental data required to obtain this model and the results collected are showcased. The behavior of the muscle when subjected to different scenarios of testing will be observed. This extensive testing will give us an understanding of how the muscle reacts in different situations. The upcoming section will discuss the different models used to describe the muscles behavior in this thesis. By using a numerical computing software, such as excel, the introduced data can be represented in graphs as plots or surfaces.

A. Test Bench:

The test bench mainly comprises of a solid frame with a fixated ruler, Figure (4). This frame has an extruded pin like shape at the top where the air muscle is hung by its end fitting. Air pipes are connected to the air muscle supplying it with compressed air from the compressor (7bar max range). Weights are used for the Force representation while spring Dynamometers measure the force developed by the muscle, the use of each depends on the test set up. A more suitable way to measure the contraction would be to use displacement sensors instead of a ruler. Also a load cell would have been more convenient. The use of these may have resulted in more detailed results with greater ease.



Figure 4. An image from the experiments carried out displaying the Experiment Test Bench.

B. Tests:

Test 1: **Isotonic** ($F=\text{constant}$):

Examining the behavior of the Contraction Ratio of an Air muscle under fixed Force and increasing Pressure. Isotonic, [7].

$$\varepsilon = f(p, F)$$

The first test requires us to find the value of the contraction ratio when the force is kept constant. The muscle is hung from the test bench as shown in Figure (5). The force is applied through weights hung at the end of the suspended air muscle. The pressure is increased by intervals of 1 bar through the air compressor. In order to find the contraction ratio, the contracted length must be measured, the use of a ruler is sufficient. We must properly assemble the air supply. once we have made sure there are no loose parts, air leakages or any safety hazards we may start testing the muscle. This contracted length is then used to find the contraction ratio, where:

Equation 3

$$\varepsilon = \frac{L_{\text{initial}} - L_{\text{contracted}}}{L_{\text{initial}}}$$

A graph of pressure verses contraction ratio is plotted using excel. A relationship is later found.

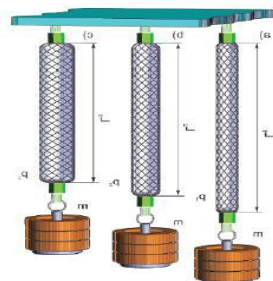


Figure 5. Isotonic test set up, $F=\text{constant}$.

Test 2: **Isometric** ($\varepsilon=\text{constant}$)

Examining the behavior of the Force developed in an Air muscle under fixed Contraction and increasing Pressure. Isometric, [7].

$$\Delta F = (p, \varepsilon)$$

The second test requires us to find the value of the Force when the Contraction ratio is kept constant. The muscle is suspended as shown in Figure (6). The contraction ratio is kept constant via weights. Once the pressure is set to the required value the Force applied due to the resistance caused by the spring dynamometer gradually pulls the muscle until it returns to the required contracted length. The hand held spring dynamometers is hung from the end of the muscle and is manually controlled. We must properly assemble the air supply. once we have made sure there are no loose parts, air leakages or any safety hazards we may start testing the muscle.

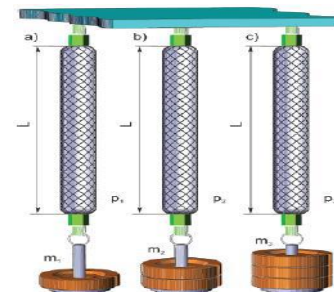


Figure 6. Isometric test set up, $\varepsilon=\text{constant}$.

As the pressure is increased the length decreases, in order to prevent this from happening we must increase the Force on the muscle. We gradually increase the load on the muscle until it returns to the contracted length we are measuring at (set point). The force is applied using a hand held spring dynamometer, the dynamometer is pulled until the muscle returns to the set point. this is the force reading to be registered.

Test 3: **Isobaric** ($p=\text{constant}$)

Examining the behavior of the Force developed in an Air muscle under fixed Pressure and increasing Contraction Ratio. Isobaric, [7].

$$F = f(\varepsilon, p)$$

The third test requires us to find the value of the Force when the Pressure is kept constant. The muscle is hung from the test bench as shown in Figure (7). The force is applied through hanging weights at the end of the muscle, while the pressure is kept constant using an air compressor. The air supply should be properly assembled, testing may proceed once we have ensured there are no loose parts, air leakages or any safety hazards.

The Pressure is increased each cycle in 1 bar intervals. The weights are added gradually and the contracted length is registered using a ruler. The contraction ratio is then found.

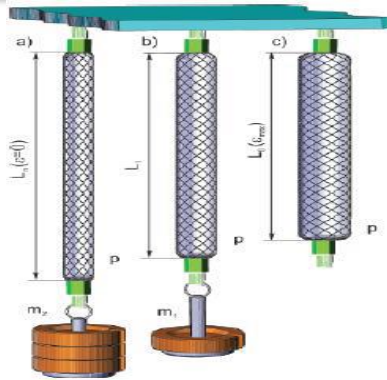


Figure 7. Isobaric test set up, $p=\text{constant}$.

• Results:

The Muscles were extensively tested until a vast amount of data was obtained. In order to interpret the data and better understand the behavior of the Muscle a plot was made to graphically represent the relations, with the use of Excel. The resulting graphs are discussed to determine whether the PAM obeyed the theoretical laws, [7].

Result of Pressure-Contraction Ratio Test:

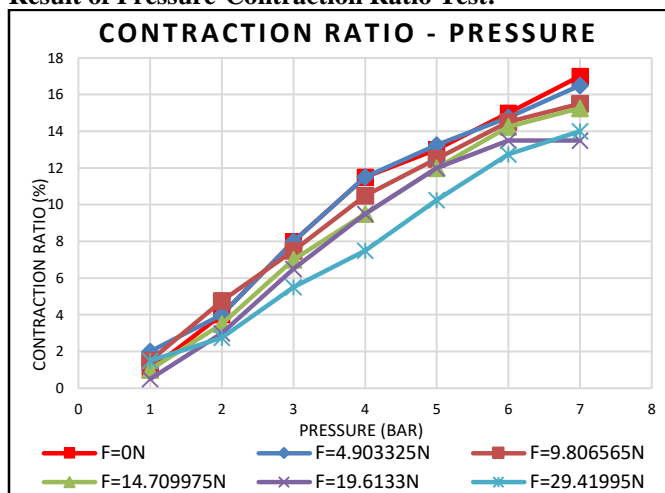


Figure 8. Test 1 Results Plot of Contraction Ratio vs. Pressure.

In Figure (8), the data collected from the pressure-contraction ratio test is plotted. Each line represents a constant force value. From the graph it is observed that, as the pressure is increased the PAM shortens to a smaller contracted length. From equation 3, it can be said that the contraction ratio increases as well, this proves Law 1. A proportional increase was expected but the results turned out to be non-linear, although they do show some form of linearity, as shown in the figure. This brings us to our next observation, the highest line in the graph is the no Force line, whereas the lowest line is the heaviest load line of 4kg. This proves Law 2.

The graph shows us that the lines are clustered together, this implies that the difference in contraction ratio is within a

small range when the force is altered, showing there is a close relationship between the force and contraction ratio.

The data proves these rules experimentally accept for a few rouge data points which could be due to a number of factors, such as sudden pressure drop, leakage and or bad data readings. It could also be an anomaly of the system. Further testing could diminish these errors.

We can conclude that, the data collected sufficiently obeys the three basic Laws of the pneumatic air muscle actuator. Seeing as the Muscle obeys these behavioral rules we can safely assume the data can be sufficiently used to obtain an Experimental model.

Result of Pressure-Force Test:

In Figure (9), the data collected from the pressure-force test is plotted. Each line represents a constant contraction ratio value. It is observed that, as the pressure is increased the pulling force developed increases, this proves Law 3.

A linear relationship was assumed although the resulting plot showed the relationship to appear slightly linear. It was observed that the force remains zero until the required fixed contraction length is obtained; only increasing once the contraction exceeds this length. The plotted line then starts to rise as the pressure increases in an almost linear fashion, with a few anomalies. The effect of the contraction ratio in this test is also considered, as the value of the contraction ratio is increased the force developed decreases, Law 2.

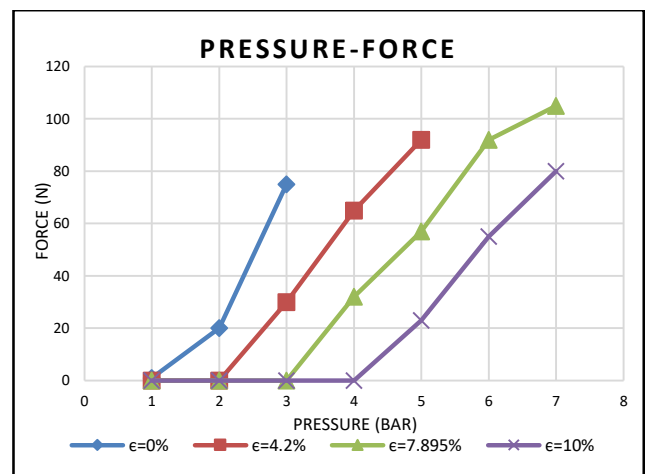


Figure 9. Test 2 Results Plot of Force vs. Pressure

From the graph we can see that it takes an extra bar for each line to start rising from its predecessor, this shows there is a fixed relationship between force and contraction even in this test scenario.

We can conclude that, the data collected follows the three basic Laws of the pneumatic air muscle actuator. Therefore, the data may be used to find an experimental model.

Result of Contraction Ratio-Force Test:

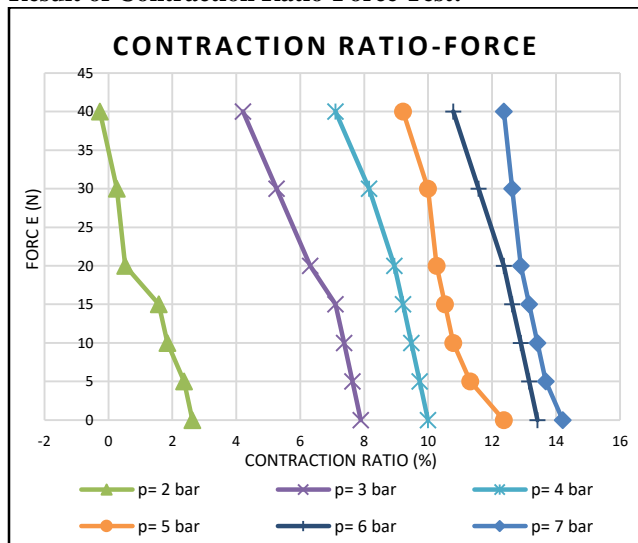


Figure 10. Test 3 Results Plot of Contraction Ratio vs. Force

In Figure (10), the data collected from the Contraction Ratio-Force test is plotted. Each line represents a constant pressure value. It is observed that, as the Contraction Ratio increases the Force decreases, Law 2.

The lines have a negative slope, ending in zero force at the largest possible contraction ratio. A linear relationship was assumed before testing. From the graph it can be seen that a non-linear relationship arises, implying our assumption was incorrect. At some points the data was close to linear. It could be observed that, a higher contraction ratio is achieved for the line which represented the highest pressure value this proves that as the pressure increases an increase in contraction ratio occurs, Law 1.

When the lines are extended to the zero contraction ratio axes then the line representing the highest pressure will correspond to the highest force again proving that as the pressure increases the force increases, Law 3.

As for the effect of the amount of pressure applied, at a pressure of 1 bar the relationship is unclear whereas there is a large gap between the contraction ratio plot of 2 bars and of 3 bars. The remaining pressures are packed closely together; this means that increasing the pressure from 2 to 3 bars will increase the contraction by a greater amount than from 3 to 4 bars, it is almost the same difference as increasing from 3 to 8 bars. Seeing as, the higher the pressure the higher the contraction ratio it can be assumed that if an even higher contraction is desired then the pressure should be increased above 7 bars. Keeping in mind the pressure limit is unknown. We can conclude that, the data collected follows the three basic rules of the pneumatic air muscle actuator. Therefore, it can be used to find an experimental model.

IV. Least Square Model

The theoretical models presented earlier are highly non linear and contain parameters that are difficult to measure or identify. To obtain a satisfactory model that can be used to design the control system we shall consider a simplified linear model. This is based on the assumption that the PAM

behavior around the operating point is linear in its physical parameters.

When using the Least Square method, a number of different combinations can be used. A simple Least Square model is used to observe the force difference that arises and whether the system obeys a linear relationship. The Force Difference is the force value calculated using the Model minus the Force experimentally observed.

A. Linear LS Model

The 54 data points were halved into two sets. Using Matlab, the first set was used to find the constants A and B whereas the second set was used to validate the resulting Linear equation. The following equation was generated:

$$F = 23.07p - 9.51\varepsilon$$

After analyzing the results, the Force Difference ranged from 26.32N to -28N and the Error Percentage range was [1548%, -826%]. We conclude that, an unacceptable Error range was generated when a Linear equation was used. Hence the Linear least square model is impractical and unreliable when calculating the Force.

The poor results obtained from the Linear LS model lead us to explore different equations. This was implemented using 3 cases. These cases are: keeping pressure linear while increasing the power of contraction ratio variable, keeping the contraction ratio linear while increasing the power of the pressure variable and increasing the power of both pressure and contraction ratio variables. The equation that resulted in the least amount of error, in comparison, was further chosen and the results displayed and discussed. It was found, based on the results, that the optimum range was obtained when both the pressure and contraction ratio were raised to the third power. Therefore, we will call this model the Least Square Model with cubed Pressure and Contraction Ratio.

B. Least Square Model with cubed Pressure and Contraction Ratio

Using Matlab, to generate the constants A_i and B_i ($i=1,2,3$), we were able to obtain the following equation from the 27 data points and verified using the remaining 27 data points.

$$F = 0.476p + 8.4754p^2 - 0.744p^3 - 5.23\varepsilon - 0.102\varepsilon^2 - 0.026\varepsilon^3.$$

The Error ranged from 100% to -1438.7%. An irrational Error percentage resulted. Whereas the difference ranged from 13.9657N in the positive direction and -17.38N in the negative direction. The difference is considered large especially since the Force is measured at low values for most data points.

The Force Difference caused by using the Cubed Pressure and Contraction Ratio LS Model is closer to zero than the Linear LS Model. This means the Cubed Pressure and Contraction Ratio LS Model is more accurate especially at a working range of 4 to 5 bars where the error Percentage is [57.3%, -21.3%], and Force difference [10.62, -3.53] (N). If a choice between the two must be made, the cubed LS model

is more acceptable. A more accurate Model is possible using other methods, as discussed in the following sections.

V. Fuzzy Logic Model

Fuzzy Logic Method is an excellent mathematical tool that handles the uncertainty that arises due to vagueness, allowing intermediate values to be defined between true/false values. These intermediate values can be formulated mathematically and processed by computers, in order to apply a more human-like way of thinking in the programming of computers. The fuzzy approach requires sufficient expert knowledge and requires a trained expert in order to operate. The employment of fuzzy logic is not recommendable if the conventional approach yields a satisfying result, an easily solvable and adequate mathematical model already exists, or the problem is not solvable. There are two types of Fuzzy Logic models, these are: Mamdani Fuzzy Inference System and Fuzzy Inference System.

A. Mamdani FLModel:

Mamdani-type FIS uses the technique of defuzzification of a fuzzy output and consists of input and output membership functions. A number of different Mamdani fuzzy logic models were prepared and analyzed with the use of Matlab, Figure (11). Each model used a unique set of handpicked rules. The resulting Force was found and compared and the error assessed. The most accurate Mamdani Fuzzy Logic Model was chosen.

At a pressure of 1 bar the system fails to reach a steady form as it struggles to overcome the resistance exerted by the elastic bladder and the friction. The bladder has yet to be properly inflated therefore it presents a non-uniform shape. we can assume the data points taken when the muscle is in this state to be irregular and do not efficiently describe the system. Excluding these data points, an Error Percentage ranging from 12.79% to -10%, a force difference range of [2.3, -4.3] N was obtained. Encouraging results using the Mamdani FIS were achieved, an improvement to the model could have been made if more data points were collected and more accurate means of measuring were used during the experiments. Air muscles are only operated within a certain working range, for this model we will set the working range between 3 to 5bars, the resulting Error Percentage is [12.79%, -9.89%] and the Force Difference (N) is [2.3, -2.7].

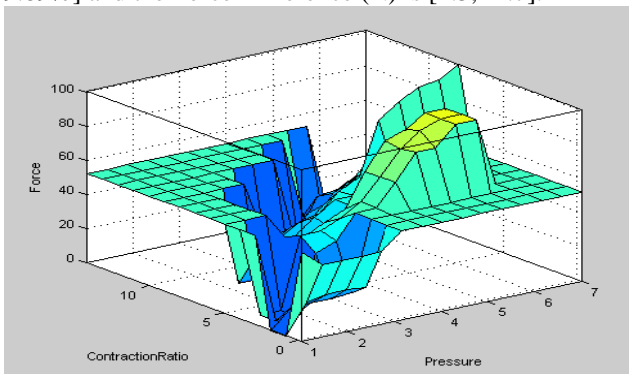


Figure 11. Surface View of Mamdani FLM

B. Sugeno FLModel:

Using the ANFIS feature in Matlab, Figure (12), we were able to train and test the system to generate a FIS. The first step to finding a model is to generate the FIS using the ANFIS feature, the training data is called upon from the workspace. The workspace contains two matrices summoned by a code. This code is comprised of two matrices each matrix holds half the 54 data points. Matrix A is used as the Training data for ANFIS and Matrix B is used as the Testing data.

Once the model had been generated, we were able to start manually inserting the input data points in the Rule Viewer, in order to find the calculated Force and compare. The Error percentage calculated ranged between 100% to -80% and the Force Difference ranged between 6.3N to -40N. This unacceptable range led us to consider a few assumptions to surpass confusion that arose due to rouge points.

- 1) At a pressure of 1 bar the data in unreliable.
- 2) The contraction ratio for one of the test points is out of range.
- 3) All data points with $F=0$ are irrelevant; in order for the Muscles to carry out their jobs a Force must be exerted.

It should be mentioned that, the fuzzy system is ill informed when it comes to large forces. The scarcity of the training data resulted in the system being unable to predict these situations. To overcome this, a broader range of testing must be carried out to increase the amount of data points retaining to these large forces for better training of the system. A working range of 3 to 5bars was set, this is reduced the error by 26.5% to -30.44%. and a Force Difference(N) range of 3.6N to -3.5N. It is clear that, very good results were obtained using the Sugeno Fuzzy Logic Model to predict the muscle Force, regardless of few anomalies.

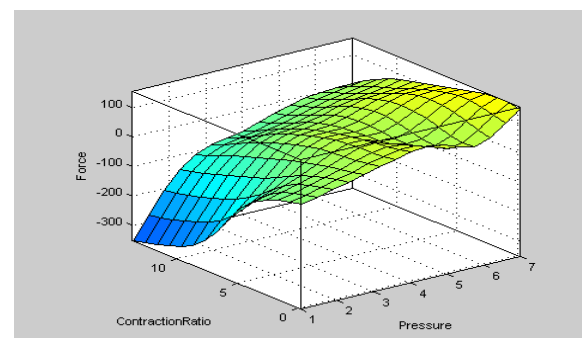


Figure 12. Surface View of Sugeno FLM

C. Comparison of Fuzzy Logic Models:

Both Fuzzy Inference System types were compared to determine which resulted in a model that best describes the Actuators' behavior. By Comparing the Error Percentage ranges and Force Difference ranges it was found that the Mamdani Fuzzy Logic Model resulted in a smaller error Percentage range and a slightly smaller Force difference range. The Mamdani FLM is more efficient than Sugeno FLM when the Force exerted by the muscle is to be found. Even though the Sugeno FIS is easier to operate and does not need an experienced analyzer to be constructed, making it more appealing. The Mamdani FLM approach requires a

more equipped individual to analyze the system and set the proper rules to achieve the best results making it less appealing.

D. Application of the model for control system simulation and design:

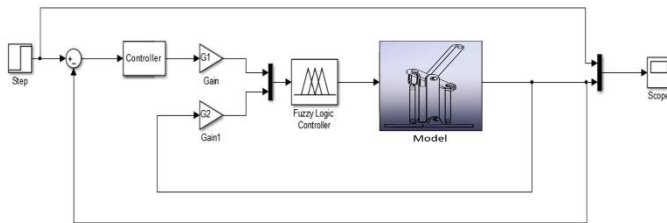


Figure 13. Control system set up

A model for the Pneumatic air muscle was successfully achieved. This model is necessary for the control of the air muscles; it represents the actuator within the control system. A control system basically comprises of: a controller, an actuator and a system. The desired value is represented by the arrow entering the feedback sum off point. The desired value is compared with the value resulting from the system, represented using a transfer function. the difference between the two values is the error. If the resulting error is zero, the controller has no job to do and the system keeps working as is. When the error is not equal to zero the controller takes in the value of the error and calculates the appropriate value for the pressure. The fuzzy logic model then uses this pressure value to calculate the resulting Force, the contraction ratio is found using sensors once the pressure is supplied. The force value is sent to the system resulting in a change in position of the joint, a sensor records the change in position and sends it back via the feedback loop. As explained if the error is zero the loop halts. If the error is found to be non-zero, then the operation is repeated until a zero error is achieved and the desired value is reached. A saturation block found in Simulink is used to assign upper and lower limits making sure the value for pressure and contraction ratio stays within our tested range.

The flow chart in Figure (13) shows the serial sequence the controller takes in making the set of decisions required to assess the required value of pressure. If the Force is too high the pressure is decreased, the vents are opened. If the Force is too low the pressure is increased, the vents are closed. If there is no error the pressure value is maintained.

I. Conclusion

In this paper, a pneumatic artificial muscle is fabricated for rehabilitation robotic system applications. In order for this muscle to be controlled effectively a reasonable model is required for control system design and simulation. The actual theoretical model is quite complicated and difficult to linearized or simplify using standard techniques. For this reason, an experimental modeling approach is considered. Models based on Least Square Method and Fuzzy Logic methods were experimentally investigated. The results obtained have shown that the FLC based models give smaller error compared to LS method.

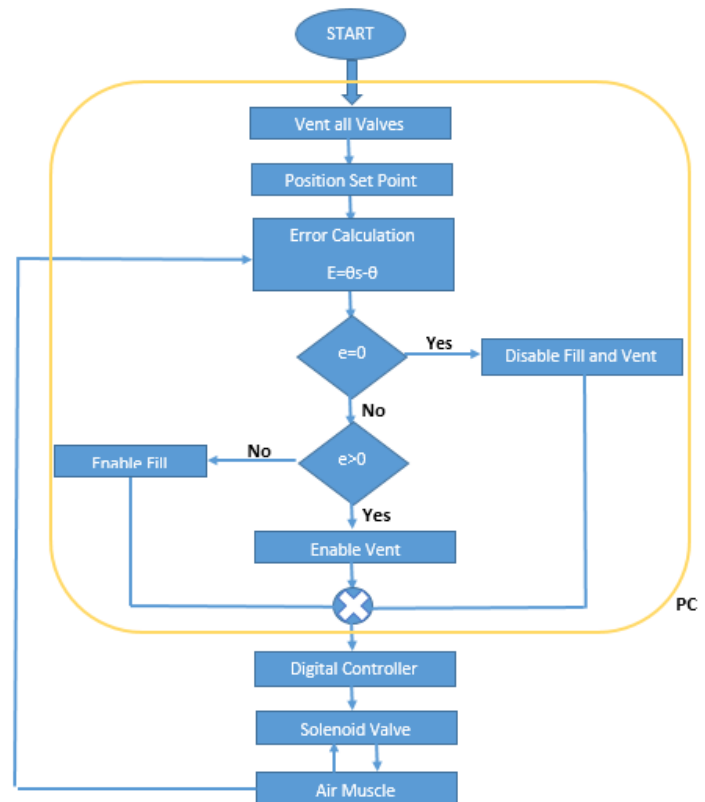


Figure 14. Flow chart for the Control process of a PAM.

References

- [1] W. Gurstelle, " ReMaking History:Makers of the World", *Volume 3, 1st Edition*, 2017.
- [2] Chavoshian, M., Taghizadeh, M. Recurrent neuro-fuzzy model of pneumatic artificial muscle position. *J Mech Sci Technol* 34, 499–508 (2020).
- [3] G. Andrikopoulos, G. Nikolakopoulos and S. Manesis, "A Survey on applications of Pneumatic Artificial Muscles," 2011 19th Mediterranean Conference on Control & Automation (MED), 2011, pp. 1439-1446, doi: 10.1109/MED.2011.5982983..
- [4] K. Mortier, "Braided Pneumatic Muscles for Rehabilitation Apparatus", 2014.
- [5] Honus, "How to Make Air Muscles!," [Online]. Available: <https://www.instructables.com/id/How-to-make-air-muscles/>.
- [6] N. Tsagarakis and D. G. Caldwell, "Improved modelling and assessment of pneumatic muscle actuators," *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, 2000, pp. 3641-3646 vol.4, doi: 10.1109/ROBOT.2000.845299..
- [7] J. E. Takosoglu, P. A. Laski, S. Blasiak, G. Bracha and D. Pietrala, "Determining the Static Characteristics of Pneumatic Muscles," 2016.
- [8] M. Hellmann, "Fuzzy Logic Introduction," in *Encyclopedia of Research Design*, 2001.
- [9] Mandal, D. Bajpai and Ahijit, "Comparative Analysis of T- Sugeno and Mamdani Type Fuzzy Logic," 2015.

How Crucial Is It for 6G Networks to Be Autonomous?

¹ Nadia Adem, ² Ahmed Benfaid, ³ Ramy Harib, and ⁴ Anas Alarabi

^{1,2,3,4} Department of Electrical and Electronic Engineering University of Tripoli,
Tripoli, Libya

¹ n.adem, ² a.benfaid@uot.edu.ly, ³ r.harib@uot.edu.ly, ⁴ an.alarabi@uot.edu.ly

ABSTRACT

The sixth generation (6G), unlike any of the previous generations, is envisioned by 2030 to connect everything. Moreover, in addition to the new use cases 6G is expected to support, it will need to provide a superior performance over 5G. The global connectivity, large network dimensions, users heterogeneity, extremely-low power consumption, high-throughput, ultrareliability, efficient network operation and maintenance, and lowlatency requirements to be met by future networks inevitably necessitate the autonomy of 6G. Intelligence, facilitated mainly by the advancement of artificial intelligence (AI) techniques, is a key to achieve autonomy. In this paper, we provide a bird's-eye view of 6G, vision, progress, and objectives. We review some technologies that would be mainly enabling the globally intelligent connected world. We, in addition to discussing the role of AI in future networks, unlike any other review papers provide our original results which give early evidence for the viability of autonomy for future wireless communications. Furthermore, we, very importantly, identify 6G implementation challenges and key innovative techniques like quantum computing and blockchain that promise to solve them. This article serves as a starting point for learners to acquire more knowledge about the field and also for researchers to promote more development.

Keywords:

I. INTRODUCTION

Recently, as the fifth generation (5G) standard has been getting finalized, and in order to set unified precise targets and roadmap for the sixth generation (6G) communication networks, researchers from industry and academia have begun to envision and extensively discuss its key values and use cases [1]–[7].

A. 6G Vision

The 6G is depicted, in its first global vision, as a framework of different services such as sensing, computing, caching, imaging, highly accurate positioning and mobility, radar and navigation integrated with the main communication services the previous generations provide. 6G will be an autonomous intelligent ecosystem that connects everything through multidimensional networks that provide services in the ground, air, space, and underwater aiming to provide quality of not just service but life. Benefiting from smart

sensory environments, 6G will be driven by a variety of verticals including

factories, automation and transportation, and healthcare. 6G is predicted to be motivated by potential-new applications for which the currently deployed 5G infrastructure is not expected to support [8]. A main domain of these applications is the connected robotics and autonomous systems that include self-driving cars, drone delivery systems, and autonomous robotics. These applications will get the full use of the multidimensional network structure, and artificial intelligence (AI) capabilities to be offered. Other domains include extended reality (XR), blockchain and wireless brain-computer interaction based applications. Different requirements for these applications and their corresponding use cases set different trends and research directions towards 6G. For example, autonomous systems applications require ultra-reliable, lowlatency, and secure communications. There are a number of applications, e.g. holographic communications

[9], that are bandwidth intensive and thus require opening up new and wider spectrum such as terahertz (THz) bands [10]. On the other hand, other use cases involve

communications between multiple small devices that mainly work on batteries, for example, set the trend for energy sustainability. More details about the requirements demanded, and technologies promising for realizing the 6G are presented in this article.

B. 6G objectives

Spectrum and energy efficiency, peak rate, user data rate, capacity per unit volume, connectivity density, latency, reliability, and mobility are some of the key performance indicators (KPIs) evaluating 6G networks. Technical objectives for these KPI, set mainly based on requirements, trends, and applications, are summarized and compared to their counterparts in 5G as follows:

- Up to 10 Tbit/s of peak data rate (1000 times that of 5G) to be targeted [2].
- A user data rate of 1 Gbit/s (10 times 5G) or higher.
- No more than 100 μ s (1=10 5G) of latency is allowed. 10 times lower latency in certain cases needs to be maintained [2].
- Providing services to 10⁷ devices/km² resulting in 10 times of the 5G connectivity density.
- *Volumetric* spectral and energy efficiency 100 times the per unit area efficiency in 5G [8].
- *Reliability* of seven 9s as opposite to that of five 9s in 5G.
- Support mobility up to 1000 kmph, as the 6G is foreseen, for example, to support airline systems and high-speed railways. The 5G, however, was intended to support less than half of that.

In this paper we discuss the key innovative technologies promising to meet some of the aforementioned objectives.

C. Paper Organization

This paper is organized as follows. Section II contains some of 6G enabling technologies, their basic principles, and other techniques they need to be integrated with to get the most out of them. In Section III, the role of AI in maintaining intelligent and hence autonomous 6G network is discussed. The Section, in addition, presents some of our simulated experiments results that demonstrate the effectiveness and superiority of AI-enabled networks. In Section IV, we cover some of future networks's challenges, possible solutions, and related open research directions. Concluding remarks are given in Section V.

II. 6G ENABLING TECHNOLOGIES

The capability expansions, performance improvements, variety of new trends and service classes that need to be guaranteed and supported by 6G require the incorporation of disruptive technologies. In this section we present some of 6G promising technologies.

A. Above 6GHz Communications

The sub-6 GHz band becomes highly congested due to heavily used frequency resources. It no longer can support massive increases in communication capacity. Allowing accessibility to higher frequencies and bandwidth, the millimeter wave (mmWave) technology has been to overcome the lack of spectrum issue. Facilitating the 6G high-rate highmobility use cases (e.g. autonomous vehicles), mmWave is considered to be one of the 6G enabling technologies [10]. Owing to its ability to offer even more abundance of bandwidth, THz is also viewed as one of the very key 6G promising technologies. MmWave communications, ranging from around 30 GHz to 300 GHz, will be one of the leading candidate systems for future wireless communications [11]. MmWave can achieve peak data rates of 10 Gbit/s or more with full-duplex capability, far exceeding the lower microwave frequency limit of 1 Gbit/s [12].

Despite its benefits, nevertheless, atmospheric absorption highly attenuates mmWave signals. Using Friis Law [13] one can determine an obtained power, P_R , at a receiver located at a distance d apart from a transmitter sending a signal with wavelength λ and power P_T as described by the following equation.

$$\frac{P_R}{P_T} = G_T G_R \left(\frac{\lambda}{4\pi d} \right)^2 \quad (1)$$

where G_R & G_T are the receiving and transmitting antenna gain respectively. The aforementioned equation works only in free space propagation. More generalized environments however are discussed in [14]. We can notice from the equation that under the same distance and antenna settings as microwave signals below 6 GHz, the free space path loss for mmWave signals is much higher. This shows that using mmWave frequencies reduces the transmission distance which is a huge disadvantage in mmWave systems. Due to their associated signals' short wavelengths, nevertheless, mmWave technology can be integrated with large-scale multielement antenna arrays allowing for high antenna gain and highly directional communications hence improving security and interference immunity. Innovative technologies like super massive (SM) multiple input multiple output (MIMO) [15] along with beamforming are considered to extend the transmission distance and hence overcoming the shortcoming of using high frequencies. Another major technical challenge, however, for mmWave is the blockage effect. Rending to their abilities in providing line-of-sight (LoS) links with a high chance, aerial base stations (BSs) and the resultant potential 6G 3D networks can decrease the impact of the aforementioned issue.

Commercial mmWave communications have become a reality with the standardization of 5G, however it has been argued that 5G's possible potentials would fall short of many new technologies, such as 3D gaming and XR [8]. Hundreds of gigabits per second to multiple terabits per second data rates are needed for these applications which will require some other innovative solutions. This leads us to THz communications systems since they are widely recognized as the next step of wireless communications research due to their ability to deliver such data rates over

short distances [10]. THz waves cover the frequency range of 0.1 to 10 terahertz. The 275GHz–3THz band spectrum is the main part of the THz band for cellular communications, according to International Telecommunication Union Radio Communication Sector (ITU-R) guidelines [16].

Despite their features, however, there are many open research issues in THz radio systems such as small propagation range due to high propagation and molecular absorption losses, and THz transceiver design that will need to be addressed to make it realizable. To achieve a successful THz deployment and operation in 6G networks, nevertheless, pencil beams [17], [18], and holographic surfaces [19] solutions are to be considered as they portend overcoming THz issues. In spite of that, the co-existence of THz systems with sub-6 GHz and mmWaves is necessary, especially in non-line-of-sight and mobility users cases, where the path loss would increase exponentially, which also gives the need of fusing machine learning (ML) techniques to make data-driven decision to get the best out of resources.

In addition to the mmWave and THz technologies like optical and visible light communication (VLC) [8] are some other above 6G non-radio frequency (RF) spectrum technologies promising for capacity boosting.

A. Non-Orthogonal Multiple Access (NOMA)

Traditional orthogonal multiple access (OMA) schemes may struggle to accommodate the massive number of 6G connections, since these techniques divide frequency resources to weak and strong users equally.

On the contrary, NOMA supports multiple users in the same time/code/frequency domain, hence promises for better spectral efficiency and throughput. Furthermore, OMA prefers users with good channel conditions which have higher schedule priority. However, in NOMA, users are served simultaneously, hence offers a better fairness, lower latency, and more connections demanded by 6G networks. To transmit two messages or more in the same time/code/frequency domains, NOMA uses superposition coding (SC) in the power domain [20], which aims to encode them into a single signal each of which gets assigned a different power value depending on intended receivers' channel conditions. Many users can be distinguished from one another at NOMA receivers by using successive interference cancellation (SIC).

NOMA, nonetheless, has some limitations, such as receiver computational complexity, BS channel state information (CSI) prior knowledge, and performance degradation as number of users increases. Due to these issues, the 3rd Generation Partnership Project (3GPP) has deferred the use of NOMA for next-generation networks, where most of these limitations will not be a concern, and given up on making NOMA an option for 5G [21], [22].

To benefit from the superiority of NOMA in fairness, latency, throughput, and connections density in the 6G networks, however, in 3D networks, NOMA will be expected to be merged with mmWave to get the best of both worlds. Optimizing such a network, nonetheless, would be a complex nonconvex problem that handling it with

traditional mathematical solutions is a challenge [23], however, we will discuss the viability of managing these networks autonomously by using the power of AI.

B. Unmanned Aerial Vehicle (UAV)

Aerial BSs, or UAVs, are emerging as one of the 6G essentials as they can provide high data rates and global wireless connectivity. UAVs enjoy exclusive features such as ease of deployment, high LoS links probability, and large degree-of-freedom provided through their controlled mobility [24]. UAVs can be deployed to offer wireless connectivity in case of emergencies such as natural disasters, overcome last mile issues, enhance capacity, replace terrestrial BSs, etc. Allowing UAVs to adaptively relocate according to environmental changes and user demands is one of the most important UAV features. The goal of connecting the unconnected to be met by the 6G will be made possible by the aerial BSs as their deployment are much easier and more economical than that of their terrestrial counterparts, especially in rural areas. Rending to their ability to connect with low earth orbit (LEO) satellites, CubeSats, and terrestrial BSs, UAVs are also envisioned to be one of the main technologies to realize the 6G multi-dimensional network. Managing UAVs' placement and resource allocation simultaneously, however, is a challenge that has to be addressed in order to get the most out of their resultant 3D networks. AI advances are promising, nonetheless, in managing the UAV networks effectively [23].

C. Quantum Communications and Computing

Quantum-assisted communication is a novel field that can be viewed as one of the cornerstones for future multi-state networks. Rendering to their capabilities in solving problems exponentially faster than their classical counterparts, quantum computers (QC) will bring a new era in telecommunication. Their working principle is based on essential concepts of quantum mechanics such as superposition, entanglement, and no-cloning theorem [25]. Constructing independent copies of quantum information is impossible, as the no-cloning theorem states. Thus, offering high communication security which should be a key feature of 6G. A qubit, quantum analogy of the classical bit, represents a two-level quantum system, where each level is called a state. Instead of representing zero or one only, however, a qubit can be in a superposition, linear combination, of both. That means upon measurement, a qubit will be found with some probability in the one or zero state. For n qubits, a quantum computer can work with 2^n quantum states simultaneously. This parallelism makes quantum computers potentially useful in applications that require the processing of big data. Many efforts have been done by the research community to investigate on the involvement of quantum speedups in various communication areas. The authors in [26] for instance have proposed and presented a number of quantum search algorithms that are particularly applicable to wireless communications. Quantum computing merged with different AI techniques, e.g. quantum machine learning (QML), has promising potentials in solving challenges in many aspects of 6G networks and beyond [27].

D. Blockchain

Blockchain is considered as a technology breakthrough in the recent years that is expected to have an important role in achieving 6G objectives [28], [29]. Blockchain, which is simply a database structure, offers multiple key characteristics that are useful when used with certain applications. Blockchain stores data in blocks called datablocks, each datablock is attached to the previous one forming a chain of blocks. This chain is replicated and stored across different nodes that form a network. Once a data update request is initiated, it is broadcasted to all nodes where they use their available computational power to verify that it follows some pre-defined rules, once verified, the datablock is linked to the previous one in the chain. The described structure along with the operation method enable the blockchain technology to be used in applications that require high level of security, where changing one datablock in a certain node can be easily detected once the other nodes cross-reference each other. Other characteristics provided by blockchain-based systems are their scalability, where they can be used for 6G to overcome scaling limitations of centralized conventional networks. Transparency is also a key feature of blockchain that can be used for dynamic resource management across different small scale operators since the records can be set to be accessed with equal rights without the need for a third party. The distributed architecture of blockchain is also beneficial in facilitating decentralized systems and thus eliminating single point failures of previous generations systems. Consequently, blockchain was proposed to be used in different areas of 6G such as blockchain-based resource management frameworks [30], and blockchain-enabled architectures for UAV networks to ensure security and privacy with enhanced network performance [31].

The aforementioned technologies promise in meeting some of the 6G objectives and enabling corresponding features including throughput, latency, connectivity, security, etc. Other technologies, whether they are evolving from the 5G or exclusive to 6G, for example large intelligent surfaces [32] and holographic beamforming [33], are out of the scope of this article.

III. INTELLIGENCE OF FUTURE NETWORKS AND AUTONOMY

In addition to communication, 6G will, exclusively, offer other services like computing, control, localization, sensing and caching for very heterogeneous use cases in a highly complex and dynamic environment. Organizing resources of 6G networks sustainably while meeting desired KPIs will inevitably need to be done autonomously. Leveraging AI advances is a must in achieving networks self organization. Rendering to the huge gain observed from incorporating AI techniques in optimizing wireless networks, research community has been giving them more and more attention [34]–[36]. Through learning and prediction, AI will be an efficient solution for achieving the

convergence in managing and allocating 6G network resources for various services [2]. Softwarization [37], cloudization [38], virtualization [39], and slicing [40] which are main techniques for 5G network orchestration, will be also expected to be important characteristics of 6G autonomous network, AI enabled however. AI algorithms will, in addition, be implemented at the network edge, for example at the smart wearable devices, enabling collective intelligence and distributed autonomy.

When it comes to optimizing resources (time, spectrum, space, beam, mode, power, and code) of multi-level and multi-dimensional networks with massiveness of connections, AI does not just outperform legacy techniques, it however emerges as a main management technique for the following reasons:

- Conventional design methods which are built based on mathematical and statistical models require perfect system characteristic knowledge which can not always be accessible.
- Traditional methods based solutions are not always optimal as their corresponding decisions are taking only according to current input and no account is taken for future information.
- Due to the complexity of a wireless network problem, mathematical solutions may not be feasible.

Rendering to their high ability in learning, predicting patterns, and taking decisions accordingly, AI techniques, on the other hand, lead to enhancing wireless network performance in situations where legacy techniques fail to [23], [41]. Managing a networks that integrates mmWave, NOMA, and UAV technologies such that the power and beam allocation, and UAV placement are jointly optimized, for example, is definitely a complicated task and traditional methods fall short to handle [23]. Using novel-AI solutions, we, nonetheless, accomplished in [23] to not only solve the problem without the need of disjointing it into sub (unrealistic) ones but also outperform some other existing heuristics derived by legacy methods e.g. [42]. We, more specifically, proposed a deep reinforcement learning (DRL) [43] based framework that simultaneously places a UAV in a 3D space and allocates NOMA power among users associated in clusters such that average sum of users' data rate is maximized and certain fairness criteria is met. In this article, we present some of the simulation results as the UAV is being trained to learn the optimal decision while serving a total of four users, where each two are associated with a certain cluster. The i^{th} user will be denoted by USER i and a decision interval will be referred to as an episode. More details about the framework and simulation setup are available in [23]. Fig. 1 shows power allocation, channel conditions, and average sum rate of our proposed framework, denoted by DRL, as they vary over episodes. In Fig. 1(c) we include the average sum rate achieved by the state of art (SoA) suggested in [42] and denoted in the figure by SoA. We can notice that the DRL average achievable sum rate converges to 23.5 Gbit/s which represents a 57% improvement compared to 15 Gbit/s

given by SoA. To be able to determine a mathematical solution for the optimization problem, the authors in [42] ended up restricting the UAV placement into a 2D plane. In contrary, our DRL framework allows 3D UAV mobility without making any confines. Hence, the gain in performance our AI based technique offers is intuitive. Furthermore, we can notice that DRL managed to recognize the NOMA power allocation users order by allocating the far user (the one with the poor channel condition as depicted in the top of Fig. 1(a) and (b)) more power (as exhibited in the bottom of Fig. 1(a) and (b)) than the near one without even imposing that in the algorithm.

The resultant choice of order is crucial for the work of SIC [20] and is the best to get the most out of NOMA.

The power allocation related decision was made intelligently and autonomously by the UAV through the careful algorithm design with the goal of improving total sum rate and maintaining a certain fairness level. The presented findings do not only show how compelling AI schemes are

in enhancing performance, but also their abilities in achieving autonomy and hence realizing future network vision.

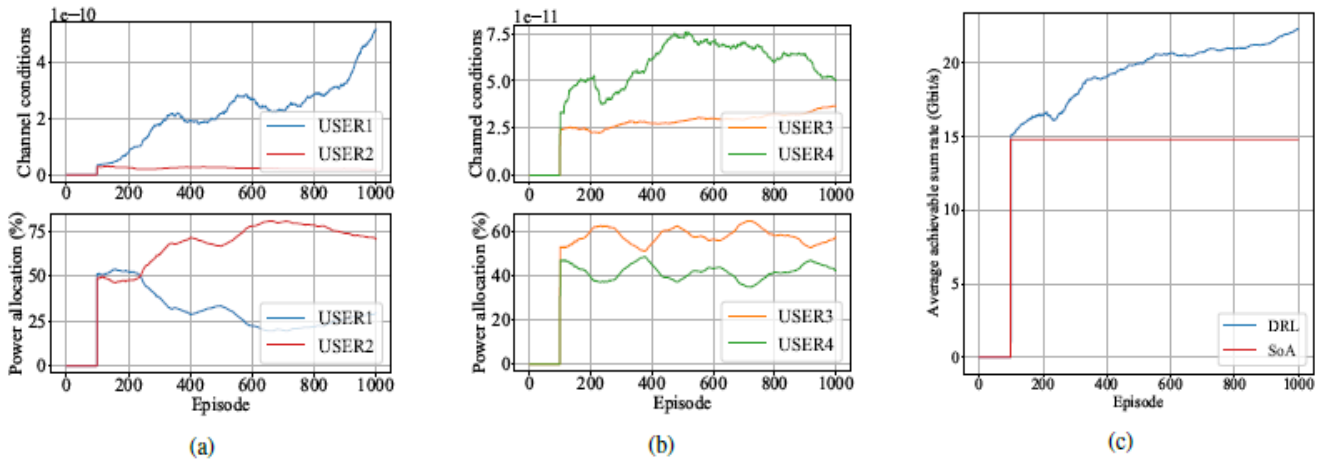


Fig. 1: DRL decision and performance without forcing NOMA order (a) & (b) cluster 1 & 2 channel conditions and power allocation respectively, and (c) total average achievable sum rate [23].

IV. FUTURE NETWORK DEPLOYMENT CHALLENGES AND SOLUTIONS

A. Complexity

The superior performance of 6G and the ability to achieve full autonomous networks requires high computation capabilities, which will be a challenging task even for most powerful computers. Despite its extremely wide spectrum, mmWave and THz bands, for example, require highly directional transmission. This means searching in a very large angular space is essential for achieving high beamforming gain and thus, reducing the initial access performance. In other words, the process of aligning beams between two communicating entities leads to communication and computational overhead. The good news, however, is that machine and deep learning-based initial access methods have the potentials in reducing such overhead [44], [45]. Moreover, the immense power of quantum computers in speeding up ML algorithms can be investigated not only for beam alignments, but also for some other services such as users tracking [27].

NOMA, even though it promises for high connectivity, it is still limited by the high computation complexity incurred for SIC estimation, especially if number of users is high, as

stated earlier. Moreover, clustering and allocating power for mobile users in mmWave-NOMA demand high computational power, as they need to be updated in real time. Emerging AI techniques can overcome some of these limitations [46].

With the ever-increasing demand for high capacity, implementing multi-user multiple input multiple output (MUMIMO) will be impractical. The utilized detection algorithms suffer from high complexity which increases exponentially with the number of users and data rate. [47] is one of the few works that proposed a solution for MU-MIMO complexity problems and showed its effectiveness by implementing it on a real quantum computer. With very low bit error rate considering different modulation schemes and channel conditions, the authors of the paper managed to achieve MU-MIMO signal detection for number of users and data rates that are unfeasible to handle in classical computers just in order of microseconds. The necessity for high computational power and parallelism suggests that quantum machine learning can be a crucial candidate in 6G networks and beyond [27].

B. Security

A number of applications including autonomous system such as healthcare, robotic, and vehicular communications demand a very-high level of security. Classical cryptographic methods would fail short with the emerging

quantum technology [48]. Through some of its attributes including anonymization, decentralization, and untraceability, blockchain technology, however, can come as a solution for security vulnerability issues [49]. UAVs for example are emerging not just as aerial BSs, but also to provide a variety of other applications such as logistics, surveillance, disaster management, and rescue operations. Securing such UAV networks which will eventually be transformed into Internet of UAVs, will inevitably require leveraging of blockchain technology. Addressing security issues for which the unique 6G network may be vulnerable to, will certainly require a lot of attention from the research community.

V. CONCLUSION

In this article we outlined the 6G vision and requirements. We presented some of its enabling technology. We also emphasized on the need of 6G network autonomy and hence intelligence. In addition, considering some of our numerical analysis, we emphasized on the necessity of integrating AI techniques in the resources management of future networks. Furthermore, the article presents some open research problems and challenges in realizing future networks objectives and trends and discussed the potentials of some emerging technologies in solving them.

References

- [1] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?" *Nature Electronics*, vol. 3, no. 1, pp. 20–29, 2020.
- [2] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019.
- [3] K. David and H. Berndt, "6G vision and requirements: Is there any need for beyond 5G?" *IEEE Vehicular Technology Magazine*, vol. 13, no. 3, pp. 72–80, 2018.
- [4] A. Yastrebova, R. Kirichek, Y. Koucheryavy, A. Borodin, and A. Koucheryavy, "Future networks 2030: Architecture & requirements," in *2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE, 2018, pp. 1–8.
- [5] F. Tariq, M. R. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 118–125, 2020.
- [6] E. Yaacoub and M.-S. Alouini, "A key 6G challenge and opportunity—connecting the remaining 4 billions: a survey on rural connectivity," *arXiv preprint arXiv:1906.11541*, 2019.
- [7] A. Pärssinen, M. Alouini, M. Berg, T. K. P. Kyösti, M. Leinonen, M. Matinmikko-Blue, E. McCune, U. Pfeiffer, and P. Wambacq, "White paper on RF enabling 6G – opportunities and challenges from technology to spectrum [white paper]," *6G Research Visions*, No. 13, 2020. [Online]. Available: <http://urn.fi/urn:isbn:9789526228419>
- [8] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.
- [9] K. Wakunami, P.-Y. Hsieh, R. Oi, T. Senoh, H. Sasaki, Y. Ichihashi, M. Okui, Y.-P. Huang, and K. Yamamoto, "Projection-type see-through holographic three-dimensional display," *Nature communications*, vol. 7, no. 1, pp. 1–7, 2016.
- [10] C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, P. Popovski, and M. Debbah, "Seven defining features of terahertz (THz) wireless systems: A fellowship of communication and sensing," *CoRR*, vol. abs/2102.07668, 2021. [Online]. Available: <https://arxiv.org/abs/2102.07668>
- [11] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, I. Chih-Lin et al., "Millimeter wave communications for future mobile networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1909–1935, 2017.
- [12] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifthgeneration (5G) wireless networks—with a focus on propagation models," *IEEE Transactions on antennas and propagation*, vol. 65, no. 12, pp. 6213–6230, 2017.
- [13] T. S. Rappaport et al., *Wireless communications: principles and practice*. prentice hall PTR New Jersey, 1996, vol. 2.
- [14] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE journal of selected topics in signal processing*, vol. 10, no. 3, pp. 436–453, 2016.
- [15] P. Yang, Y. Xiao, M. Xiao, and S. Li, "6G wireless communications: Vision and potential techniques," *IEEE Network*, vol. 33, no. 4, pp. 70–75, 2019.
- [16] R.-A. Stoica and G. T. F. de Abreu, "6G: the wireless communications network for collaborative and AI applications," *arXiv preprint arXiv:1904.03413*, 2019.
- [17] J. Tan and L. Dai, "Thz precoding for 6g: Applications, challenges, solutions, and opportunities," *arXiv preprint arXiv:2005.10752*, 2020.
- [18] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *Ieee Access*, vol. 7, pp. 78 729–78 757, 2019.

- [19] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. Di Renzo, and M. Debbah, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Communications*, vol. 27, no. 5, pp. 118–125, 2020.
- [20] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181–2195, 2017.
- [21] 3GPP, "Study on non-orthogonal multiple access (NOMA) for NR," 3rd Generation Partnership Project (3GPP), Technical report (TR) 38.812, Dec. 2018. [22] B. Makki, K. Chitti, A. Behravan, and M.-S. Alouini, "A survey of NOMA: Current status and open research challenges," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 179–189, 2020.
- [22] A. Benfaid, "UAV communications and AI technologies: Pro gamers in the 6G sky," B.S. Thesis, Elect. and Electron. Eng. Dept., Univ. of Tripoli, Tripoli, Libya, 2021.
- [23] B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5G and beyond: Recent advances and future trends," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2241–2263, 2018.
- [24] L. Gyongyosi and S. Imre, "A survey on quantum computing technology," *Computer Science Review*, vol. 31, pp. 51–71, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1574013718301709>
- [25] P. Botsinis, D. Alanis, Z. Babar, H. V. Nguyen, D. Chandra, S. X. Ng, and L. Hanzo, "Quantum search algorithms for wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1209–1242, 2019.
- [26] S. J. Nawaz, S. K. Sharma, S. Wyne, M. N. Patwary, and M. Asaduzzaman, "Quantum machine learning for 6G communication networks: State-of-the-art and vision for the future," *IEEE Access*, vol. 7, pp. 46 317–46 350, 2019.
- [27] T. Hewa, G. Gūr, A. Kalla, M. Ylianttila, A. Bracken, and M. Liyanage, "The role of blockchain in 6G: Challenges, opportunities and research directions," in *2020 2nd 6G Wireless Summit (6G SUMMIT)*, 2020, pp. 1–5.
- [28] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "Blockchain for 5G and beyond networks: A state of the art survey," *Journal of Network and Computer Applications*, vol. 166, p. 102693, 2020.
- [29] H. Xu, P. V. Klaine, O. Onireti, B. Cao, M. Imran, and L. Zhang, "Blockchain-enabled resource management and sharing for 6G communications," *Digital Communications and Networks*, vol. 6, no. 3, pp. 261–269, 2020.
- [30] R. Gupta, A. Nair, S. Tanwar, and N. Kumar, "Blockchain-assisted secure UAV communication in 6G environment: Architecture, opportunities, and challenges," *IET Communications*, vol. 15, no. 10, pp. 1352–1367, 2021.
- [31] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8238–8242, 2019.
- [32] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. Di Renzo, and M. Debbah, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Communications*, vol. 27, no. 5, pp. 118–125, 2020.
- [33] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Communications surveys & tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019.
- [34] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2595–2621, 2018.
- [35] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2016.
- [36] I. Afolabi, T. Taleb, K. Samdanis, A. Ksentini, and H. Flinck, "Network slicing and softwarization: A survey on principles, enabling technologies, and solutions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2429–2453, 2018.
- [37] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks—a technology overview," *IEEE Communications surveys & tutorials*, vol. 17, no. 1, pp. 405–426, 2014.
- [38] R. Mijumbi, J. Serrat, J.-L. Gorricho, S. Latr'e, M. Charalambides, and D. Lopez, "Management and orchestration challenges in network functions virtualization," *IEEE Communications Magazine*, vol. 54, no. 1, pp. 98–105, 2016.
- [39] H. Zhang, N. Liu, X. Chu, K. Long, A.-H. Aghvami, and V. C. Leung, "Network slicing based 5G and future mobile networks: mobility, resource management, and challenges," *IEEE communications magazine*, vol. 55, no. 8, pp. 138–145, 2017.
- [40] N. Adem and B. Hamdaoui, "Jamming resiliency and mobility management in cognitive communication networks," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–6.

- [41] X. Liu, J. Wang, N. Zhao, Y. Chen, S. Zhang, Z. Ding, and F. R. Yu, "Placement and power allocation for NOMA-UAV networks," IEEE Wireless Communications Letters, vol. 8, no. 3, pp. 965–968, 2019.
- [42] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," arXiv preprint arXiv:1312.5602, 2013.
- [43] A. Mazin, M. Elkourdi, and R. D. Gitlin, "Accelerating beam sweeping in mmWave standalone 5G new radios using recurrent neural networks," in 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), 2018, pp. 1–4.
- [44] N. J. Myers, Y. Wang, N. Gonz´alez-Prelcic, and R. W. Heath, "Deep learning-based beam alignment in mmWave vehicular networks," in ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 8569–8573.
- [45] J. Cui, Z. Ding, and P. Fan, "The application of machine learning in mmWave-NOMA systems," in 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), 2018, pp. 1–6.
- [46] M. Kim, D. Venturelli, and K. Jamieson, "Leveraging quantum annealing for large MIMO processing in centralized radio access networks," in Proceedings of the ACM Special Interest Group on Data Communication, 2019, pp. 241–255.
- [47] L. Chen, L. Chen, S. Jordan, Y.-K. Liu, D. Moody, R. Peralta, R. Perlner, and D. Smith-Tone, Report on post-quantum cryptography. US Department of Commerce, National Institute of Standards and Technology, 2016, vol. 12.
- [48] R. Henry, A. Herzberg, and A. Kate, "Blockchain access privacy: Challenges and directions," IEEE Security & Privacy, vol. 16, no. 4, pp. 38–45, 2018.