

Comparative analysis of influencing factors on pedestrian road accidents

Farrukh Hafeez¹, Usman Ullah Sheikh¹, Saud Al-Shammari², Muhammad Hamid³, Abdul Baqi Khan Khakwani², Zeeshan Ahmad Arfeen⁴

¹School of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

²Department of Electrical and Electronic Engineering Technology, Jubail Industrial College, Al Jubail, Saudi Arabia

³Department of Engineering Technology, Community College of Qatar, Doha, Qatar

⁴Department of Electrical Engineering, The Islamia University of Bahawalpur (IUB), Bahawalpur, Pakistan

Article Info

Article history:

Received Apr 30, 2022

Revised Aug 12, 2022

Accepted Sep 4, 2022

Keywords:

Comparative analysis

Influencing factors

Pedestrians

Road accidents

Safety

ABSTRACT

Road accident data includes detailed information about incidents that occurred, such as where they happened, the severity of the accident, and the number of people on the road at the time. Such information is useful in determining the causes of accidents and developing potential countermeasures. This research aims to determine the factors that contribute to pedestrian fatalities and injuries in traffic accidents. This study examined 150 pedestrian-vehicle accidents that took place between 1990 and 2021 in forty countries. Eleven factors have been identified as the major causes of accidents. The categorical principal component analysis (CATPCA) technique is used to reduce the number of dimensions and identify the elements that contribute to accidents. The eleven variables are classified into three groups: human factors, roadway environment, and vehicle attributes. The study found that car speed, weather, lighting, traffic conditions, area types, accident locations, and road conditions all had a significant impact on pedestrian accidents and fatalities. The findings show that a pedestrian's state (walking, running) and intention significantly increase the risk of serious injuries and death. The analysis of the driver's status suggests that the driver's intentions may also play a role in car accidents.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Farrukh Hafeez

School of Electrical Engineering, Universiti Teknologi Malaysia

UTM Johor Bahru, 81310 Johor, Malaysia

Email: hafeez@graduate.utm.my

1. INTRODUCTION

The number of car accidents has gone up around the world, making road safety a major concern. According to the World Health Organization (WHO), over 1.35 million people die annually in road traffic accidents worldwide [1]. The injuries sustained in these crashes are considered to be the eighth largest cause of death worldwide for all age groups, implying that more individuals are being killed in road traffic accidents. Worldwide, pedestrian deaths accounted for 22% of total road user deaths in 2019 [2]. Pedestrians are more likely to have an accident or be severely injured in road accidents because they are exposed to the road environment and are less protected than vehicle occupants. Previous studies have found that factors like alcohol consumption, age, gender of drivers, pedestrian behavior, crash location, vehicle, speed, and adverse weather all affect the severity of pedestrian injuries [3]–[5]. Because the road environment and traffic are different in developing countries and industrialized countries, the things that could make pedestrian injuries worse in developing countries [6] might be different from those in industrialized countries. Poor planning and

inconsistent road design, lack of vehicle traffic, lack of safe crossing areas, and high motor speeds are all things that can cause injuries and deaths to pedestrians [7]. Previous studies revealed that pedestrians who were unconscious, drivers who had a negative attitude, vehicles that had problems, and weak traffic control and management system were all contributors to the crashes.

It is critical to identify and characterize the risk factors that contribute to the severity of injuries sustained in pedestrian-vehicle accidents. This is important for figuring out what to do, and it could help traffic engineers, planners, and decision-makers figure out what factors are causing the problem when they design solutions. However, several critical questions about pedestrian safety remain unanswered. Since pedestrians will likely get hurt badly in car accidents, it's important to study many different risk factors for pedestrian crash injury severity that haven't been looked into yet.

Thus, understanding the relationship between these risk factors and the severity of the injury will create the framework for designing safety countermeasures against pedestrian collisions and for building a walkable environment. The current study used pedestrian accidents to construct hierarchical models to assess various risk factors for pedestrian injury. The goals of this study are: i) to find out what causes pedestrian accidents so that pedestrians and vehicles can interact safely; ii) to find out how pedestrian gender, age, intention, state, and crossbehavior are related; and iii) to find out how location, time, weather, and context affect pedestrian accidents. The remainder of this paper will be structured as follows: section 2 summarises relevant work, while section 3 details the method. Section 4 contains the results, which are interpreted in light of the data analysis and objectives. The discussion and results of the study are in section 5, and the study's conclusion is in section 6.

2. RELATED WORK

In recent years, traffic safety researchers have paid a lot of attention to the analysis of crashes involving pedestrians. Identifying factors that contributed to injury severity and predicting pedestrian-vehicle accidents were the most studied parts. Many different aspects of human behavior [8], as well as road and environmental design, have been observed to severely influence the frequency and severity of pedestrian-vehicle collisions.

Factors influencing pedestrian accidents can be generally categorized as traffic and roadway characteristics, environmental factors, temporal characteristics, vehicle attributes, driver characteristics, and pedestrian characteristics [9]. Traffic patterns, roadway layouts, and environmental factors significantly influence the severity of pedestrian fatalities. Poor weather and inadequate lighting often lead to more severe injuries on the roads [10]. A significant number of previous studies employ multiple interpretations of severity and attempt to specifically identify seriously injured people in their findings. Wet roads, rain, and low light all make it more likely that a pedestrian will get hurt seriously. Previous research has shown that the risk of death for pedestrians is generally higher for men than for women, with some exceptions for people over 75 years old [5]. Even in the older age groups, the risk is higher for men than for women when population size is taken into account. Specifically, younger (25 years) and elderly (>75 years) pedestrians are at increased risk.

Among the pedestrian factors, pedestrian age and pedestrian crossing features had the greatest influence on injury results [11]. Pedestrian accidents are more likely to occur when pedestrians are under the age of 18, while pedestrians over the age of 65 are more likely to die in accidents [12]. Young pedestrians (children and teens) and elderly pedestrians (over 55) are more likely to be injured, whereas adults and middle-aged pedestrians might be severely injured. Pedestrian accidents are more common in metropolitan areas than in rural ones [13]. In terms of accident severity, fatalities are more common in rural areas and areas with higher speed limits, mostly because of the link between speed and the severity of pedestrian injuries. Previous studies have also revealed that the severity of pedestrian accidents increases with heavy vehicles. Schools, bus stops, urban sites, dry weather, local streets, spring season, industrial and unoccupied land, pedestrian injury severity is likely to be reduced [14], while split and multi-lane highways, midblock pedestrian crossings, two-way divided roads, visual obstacles, sidewalk presence, lack of traffic control devices, uncontrolled mid-block, curve segments, and sloping roads have all been shown to increase the likelihood of injury outcomes [14].

Accidents are likely to happen where there are more right-turn-only lanes and more non-residential driveways within 15 meters of an intersection. The time of day is also one of the most important factors in figuring out if a pedestrian injury will get worse, period from midnight to 6 am leads to the rising number of deaths. It was also found that night-time collisions (between 20:00 and 05:00) doubled the risk of fatal injury. Another research [15] is covered that 10 am to 3:59 pm significantly reduces the chance of fatal injuries. Some researchers have used AI-based models for the early detection of intuition [16], [17]. Some studies have shown that heavier cars increase the chance of pedestrian fatalities [18]. Motorcyclists and other people

who travel in light or compact vehicles (such as SUVs and automobiles) have a lower death rate. It has been seen that both the way cars turn and unsafe vehicles make it more likely that someone will get seriously hurt. Most of the studies have been done, and each one is based on the social, economic, and cultural conditions of the country where it was proposed.

Different strategies have been applied to know the factors that cause pedestrian accidents, like survey-based questionnaires, interviews with the victims of accidents, and modeling-based approaches. But it was found that most specific factors like gender, age [13], [19], and urban areas have been studied, while rural areas have been linked with the factors of over speeding and poor infrastructure [20]. But in the present study, we are considering both human factors and vehicle-related factors that cause pedestrian road accidents while taking into account different surrounding contexts, geographical distribution, different weather conditions, daytimes, and other factors.

3. METHOD

Numerous frameworks have been used to determine the components that contribute to road accidents. The paper provides a framework in Figure 1 that includes all of these impacting factors into a single model. Eleven variables are taken into account and chosen for this current study. The variables gender, age, pedestrian state, and driver state, as well as pedestrian intention, are grouped under the heading 'human characteristic'. Variables relating to the road environment, such as atmospheric conditions (rain, fog, snow), road markings, location, and lighting condition, are grouped in the 'roadway feature' group; variables relating to vehicles, such as vehicle speed and traffic volume, are grouped in 'vehicle properties' group. The primary data source used in this study is the international traffic safety data and analysis group (IRTAD) database [21]. The database encompasses accident records spanning over 32 years from 1990 to 2021 in 42 countries.

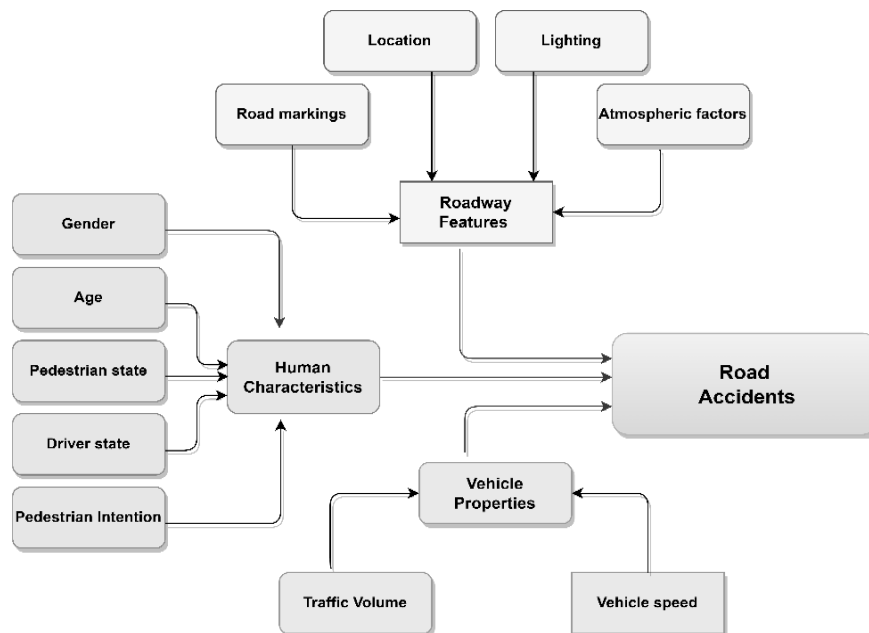


Figure 1. Factors involved in traffic accidents

3.1. Comprehensive analysis of influencing factors

Categorical principal component analysis (CATPCA) is used to reduce the number of factors in the data analysis since the variables are categorized. The CATPCA analysis in statistical program for social science (SPSS) begins with the first iteration and provides the iteration history, model summary, and variables accounted for. In contrast to typical PCA, CATPCA starts with a large eigenvalue and increases iteratively. As eigenvalues are used to calculate the percentage of variance explained (a subset of effect size), greater eigenvalues are preferred over lower ones. The model summary provides the internal consistency coefficient (Cronbach's alpha) for each dimension, its eigenvalue, and the variance imposed by each dimension. Each iteration should increase in Cronbach's alpha' value. A total of 11 variables are classified into three groups and subjected to CATPCA analysis using SPSS in three iterations. After reaching the

convergence test value in the first iteration, the operation was stopped at iteration 16. In contrast to conventional PCA, CATPCA analysis begins with a large eigenvalue and gradually improves it through iterations. In this regard, the eigenvalue was 5.2494 at iteration zero and increased to 5.6428 after the sixteenth iteration, as shown in Table 1. Similarly, the values of the centroid coordinates drop over time, showing that variables close in on the components. As shown in Table 1, the eigenvalue increases from the first to the final ideal iteration.

Table 1. Iteration history

| Iteration no. | Variance accounted for | | Total | Loss centroid coordinates | Restriction of centroid to vector coordinates |
|---------------|------------------------|----------|-----------|---------------------------|---|
| | Total | Increase | | | |
| 0 | 5.249409 | 0.000419 | 30.750591 | 29.773164 | 0.977428 |
| 1 | 5.487821 | 0.238412 | 30.512179 | 29.773164 | 0.739015 |
| 2 | 5.610466 | 0.122645 | 30.389534 | 29.583473 | 0.806062 |
| 3 | 5.631491 | 0.021025 | 30.368509 | 29.557096 | 0.811414 |
| 4 | 5.638110 | 0.006619 | 30.361890 | 29.552456 | 0.809434 |
| 5 | 5.641223 | 0.003113 | 30.358777 | 29.550365 | 0.808412 |
| 6 | 5.642173 | 0.000950 | 30.357827 | 29.549091 | 0.808736 |
| 7 | 5.642448 | 0.000275 | 30.357552 | 29.548266 | 0.809286 |
| 8 | 5.642586 | 0.000137 | 30.357414 | 29.547795 | 0.809619 |
| 9 | 5.642665 | 0.000080 | 30.357335 | 29.547516 | 0.809819 |
| 10 | 5.642717 | 0.000051 | 30.357283 | 29.547343 | 0.809941 |
| 11 | 5.642752 | 0.000035 | 30.357248 | 29.547231 | 0.810017 |
| 12 | 5.642777 | 0.000025 | 30.357223 | 29.547157 | 0.810066 |
| 13 | 5.642796 | 0.000019 | 30.357204 | 29.547106 | 0.810098 |
| 14 | 5.642810 | 0.000014 | 30.357190 | 29.547069 | 0.810121 |
| 15 | 5.642822 | 0.000011 | 30.357178 | 29.547042 | 0.810136 |
| 16 | 5.642830 | 0.000009 | 30.357170 | 29.547021 | 0.810148 |

Eleven items are explained in Table 2, approximately 47.024% of the scale variation, divided into three factors. 19.223% of the variance is accounted for by the first dimension, 14.132% by the second dimension, and 13.668% by the third dimension. The total model variance in the optimally scaled items has a Cronbach alpha of 0.898, which is near excellent. Cronbach's alpha measures internal consistency and shows how closely related a set of items is as a group. The alpha for the first component is relatively high (0.618) compared to other dimensions, indicating that items in this component are highly related compared to the other two dimensions, as indicated in Table 2.

Table 2. Model summary of the second iteration

| Dimension | Cronbach's alpha | Variance accounted for | |
|-----------|------------------|------------------------|---------------|
| | | Total (eigenvalue) | % of variance |
| 1 | 0.618 | 2.307 | 19.223 |
| 2 | 0.448 | 1.696 | 14.132 |
| 3 | 0.426 | 1.640 | 13.668 |
| Total | 0.898 | 5.643 | 47.024 |

Table 3 presents the variances accounted for by each variable. The main focus in the given table is the mean centroid coordinates. It displays the coordinates for each item on each dimension about the centroid (0, 0) when all the items are represented by a straight line between dimension 1 (x-axis) and dimension 2 (y-axis). The mean is below 0.100, which is considered low, and the variable ought to be removed. In this case, gender and traffic volume have a centroid mean value of 0.94 each and thus are removed for the second iteration to be performed. Table 3 also indicates the percentage of variance accounted for by each dimension and the eigenvalue, which equates to Table 2 results. The variance accounted for by each dimension decreases from the first one to the third, indicating that items in the first component highlighted the cause of traffic incidence.

The second iteration shows that removing gender and traffic volume increased the value of Cronbach's alpha to 0.908, and the overall percentage of variance increased to 54.816, as indicated in Table 4. The percentage of variance explained by each component increased, with the first component recording the highest increase. Similarly, the alpha for the first dimension increased to 0.627, indicating a higher association of items in this dimension. However, Cronbach's alpha for the third dimension reduced from 0.426 to 0.366, indicating a decline in inconsistency among items in this dimension.

Table 3. Variance accounted for table of first iteration

| Variables | Centroid coordinates | | | | Total (vector coordinates) | | | |
|----------------------|----------------------|-------|-------|-------|----------------------------|--------|--------|--------|
| | Dimension | | | Mean | Dimension | | | Total |
| | 1 | 2 | 3 | | 1 | 2 | 3 | |
| Gender | 0.003 | 0.001 | 0.276 | 0.094 | 0.003 | 0.001 | 0.276 | 0.281 |
| Age | 0.011 | 0.047 | 0.284 | 0.114 | 0.003 | 0.037 | 0.266 | 0.306 |
| Pedestrian state | 0.041 | 0.684 | 0.018 | 0.248 | 0.040 | 0.676 | 0.004 | 0.720 |
| Pedestrian intention | 0.217 | 0.659 | 0.153 | 0.343 | 0.062 | 0.586 | 0.123 | 0.770 |
| Location | 0.495 | 0.147 | 0.033 | 0.225 | 0.471 | 0.118 | 0.004 | 0.592 |
| Lighting | 0.099 | 0.143 | 0.188 | 0.143 | 0.026 | 0.115 | 0.152 | 0.293 |
| Traffic volume | 0.036 | 0.045 | 0.201 | 0.094 | 0.001 | 0.045 | 0.201 | 0.247 |
| Vehicle speed | 0.357 | 0.032 | 0.211 | 0.200 | 0.357 | 0.032 | 0.211 | 0.600 |
| Driver state | 0.082 | 0.042 | 0.229 | 0.118 | 0.021 | 0.019 | 0.219 | 0.259 |
| Road markings | 0.471 | 0.032 | 0.019 | 0.174 | 0.461 | 0.027 | 0.013 | 0.502 |
| Atmospheric factors | 0.373 | 0.066 | 0.134 | 0.191 | 0.336 | 0.023 | 0.106 | 0.465 |
| Active total | 2.712 | 1.924 | 1.817 | 2.151 | 2.307 | 1.696 | 1.640 | 5.643 |
| % of variance | | | | | 22.599 | 16.035 | 15.141 | 47.024 |

Table 4. Model summary of the second iteration

| Dimension | Cronbach's alpha | Variance accounted for | |
|-----------|------------------|------------------------|---------------|
| | | Total (eigenvalue) | % of variance |
| 1 | 0.627 | 2.295 | 22.948 |
| 2 | 0.455 | 1.695 | 16.949 |
| 3 | 0.366 | 1.492 | 14.918 |
| Total | 0.908 | 5.482 | 54.816 |

Table 5 presents variance accounted for Table 4, which showed that all the items had a mean value greater than 0.1 except age, which was exactly equal to 0.1. The contribution of this item was potentially questionable and thus was eliminated from the construct. A third iteration was then performed to assess whether the internal consistency would improve and the variance accounted for by the factors.

Table 5. Variance accounted for table of second iteration

| Variables | Centroid coordinates | | | | Total (vector coordinates) | | | |
|----------------------|----------------------|--------|--------|--------|----------------------------|--------|--------|--------|
| | Dimension | | | Mean | Dimension | | | Total |
| | 1 | 2 | 3 | | 1 | 2 | 3 | |
| Age | 0.011 | 0.036 | 0.254 | 0.100 | 0.004 | 0.018 | 0.251 | 0.272 |
| Pedestrian state | 0.034 | 0.688 | 0.032 | 0.251 | 0.034 | 0.676 | 0.029 | 0.739 |
| Pedestrian intention | 0.199 | 0.729 | 0.097 | 0.342 | 0.035 | 0.670 | 0.048 | 0.753 |
| Location | 0.506 | 0.117 | 0.021 | 0.215 | 0.487 | 0.085 | 0.019 | 0.590 |
| Lighting | 0.096 | 0.131 | 0.367 | 0.198 | 0.026 | 0.096 | 0.344 | 0.466 |
| Vehicle speed | 0.373 | 0.044 | 0.182 | 0.199 | 0.373 | 0.044 | 0.182 | 0.598 |
| Driver state | 0.079 | 0.070 | 0.392 | 0.180 | 0.018 | 0.039 | 0.386 | 0.443 |
| Road markings | 0.470 | 0.039 | 0.048 | 0.186 | 0.461 | 0.035 | 0.042 | 0.539 |
| Atmospheric factors | 0.353 | 0.055 | 0.147 | 0.185 | 0.313 | 0.014 | 0.118 | 0.445 |
| Active total | 2.666 | 1.939 | 1.625 | 2.077 | 2.295 | 1.695 | 1.492 | 5.482 |
| % of variance | 26.657 | 19.394 | 16.255 | 20.769 | 22.948 | 16.949 | 14.918 | 54.816 |

The third iteration aimed to eliminate the age variable. After removing all volumes from the construct, the percentage of variance increases to 59.896 with a Cronbach's alpha of 0.916, as indicated in Table 6. This suggests the excellent reliability of the CATPCA model. Although the overall alpha increased and the other first two dimensions, Cronbach's alpha for the third component declined, further, to 0.323 from 0.366. The decrease indicated a further loss of inconsistency among its items, as indicated in Table 6.

Table 6. Model summary of the second iteration

| Dimension | Cronbach's alpha | Variance accounted for | |
|-----------|------------------|------------------------|---------------|
| | | Total (eigenvalue) | % of variance |
| 1 | 0.633 | 2.285 | 25.386 |
| 2 | 0.464 | 1.703 | 18.924 |
| 3 | 0.323 | 1.403 | 15.587 |
| Total | 0.916a | 5.391 | 59.896 |

According to the variance accounted for in Table 7, none of the items has a small mean centroid coordinate value, suggesting that all the items may have suitable contributions to the model. Hence, no further iteration is required. Tables also showed that location, vehicle speed, road markings, and weather conditions affected the first dimension; pedestrian state and pedestrian intentions affected the second dimension; and lighting and driver state had a big effect on the third dimension.

Table 7. Variance accounted for table of third iteration

| Variables | Centroid coordinates | | | | Total (vector coordinates) | | | |
|----------------------|----------------------|--------|--------|--------|----------------------------|--------|--------|--------|
| | Dimension | | | Mean | Dimension | | | Total |
| | 1 | 2 | 3 | | 1 | 2 | 3 | |
| Pedestrian state | 0.025 | 0.672 | 0.022 | 0.240 | 0.025 | 0.661 | 0.022 | 0.708 |
| Pedestrian intention | 0.178 | 0.774 | 0.043 | 0.332 | 0.010 | 0.737 | 0.007 | 0.754 |
| Location | 0.516 | 0.105 | 0.050 | 0.224 | 0.500 | 0.068 | 0.050 | 0.619 |
| Lighting | 0.097 | 0.123 | 0.448 | 0.222 | 0.026 | 0.086 | 0.434 | 0.546 |
| Vehicle speed | 0.400 | 0.041 | 0.147 | 0.196 | 0.400 | 0.041 | 0.147 | 0.588 |
| Driver state | 0.073 | 0.082 | 0.443 | 0.199 | 0.014 | 0.054 | 0.426 | 0.493 |
| Road markings | 0.456 | 0.048 | 0.085 | 0.196 | 0.446 | 0.041 | 0.074 | 0.561 |
| Atmospheric factors | 0.352 | 0.045 | 0.156 | 0.184 | 0.316 | 0.003 | 0.126 | 0.445 |
| Active total | 2.646 | 1.911 | 1.520 | 2.025 | 2.285 | 1.703 | 1.403 | 5.391 |
| % of variance | 29.397 | 21.228 | 16.884 | 22.503 | 25.386 | 18.924 | 15.587 | 59.896 |

4. RESULT

The bar graph in Figure 2 is derived from Table 7 to depict the contribution of each variable to the dimension virtually. The location has the highest contribution, followed by road markings, vehicle speed, and atmospheric factors. The variables could be used to explain the underlying factor in the first component.

The contribution of each variable to the variance of the second component is depicted in Figure 3. The intention of pedestrians contributes one of the most variances in this scenario, followed by pedestrian states. As illustrated by the chart, other variables have a negligible effect. So, the two variables could be used to figure out what the underlying factor is in the second dimension.

Lighting and driver state contribute significantly more variance to the third dimension. Other variables, such as vehicle speed, appeared to have a minor effect, as do atmospheric elements, as illustrated in the chart in Figure 4. In Figure 5, the scatterplot matrix displays the object scores for each case. These object scores are three-dimensional coordinates connected with each case. The chart illustrates that the majority of cases are located and distributed consistently around the centroid (0, 0), with no extreme or outlier.

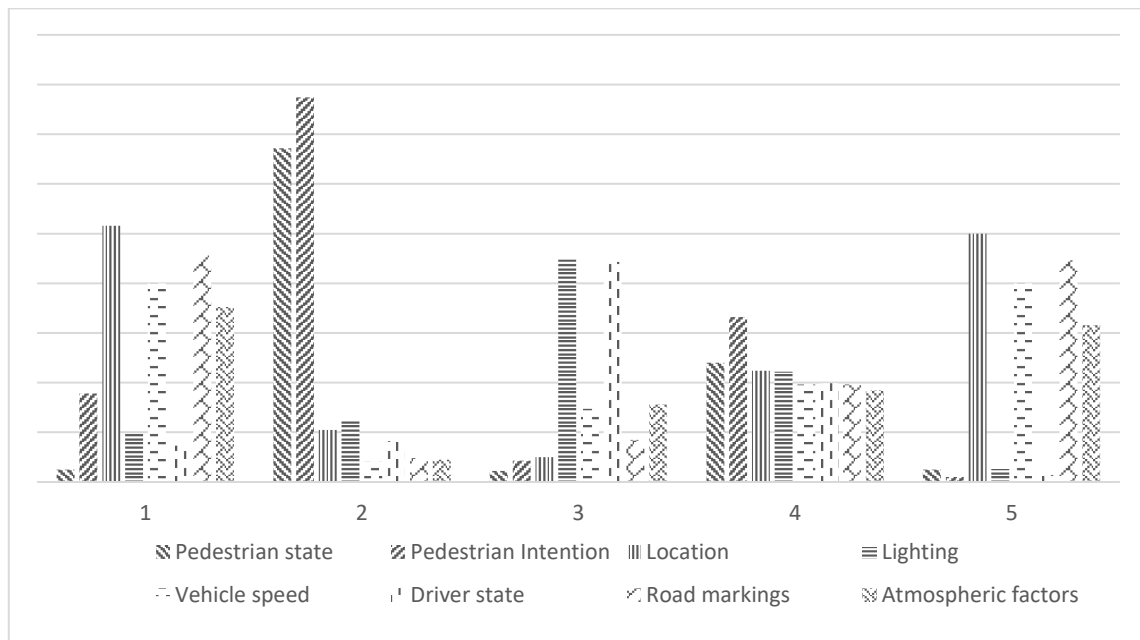


Figure 2. Variables contribution to dimension 1

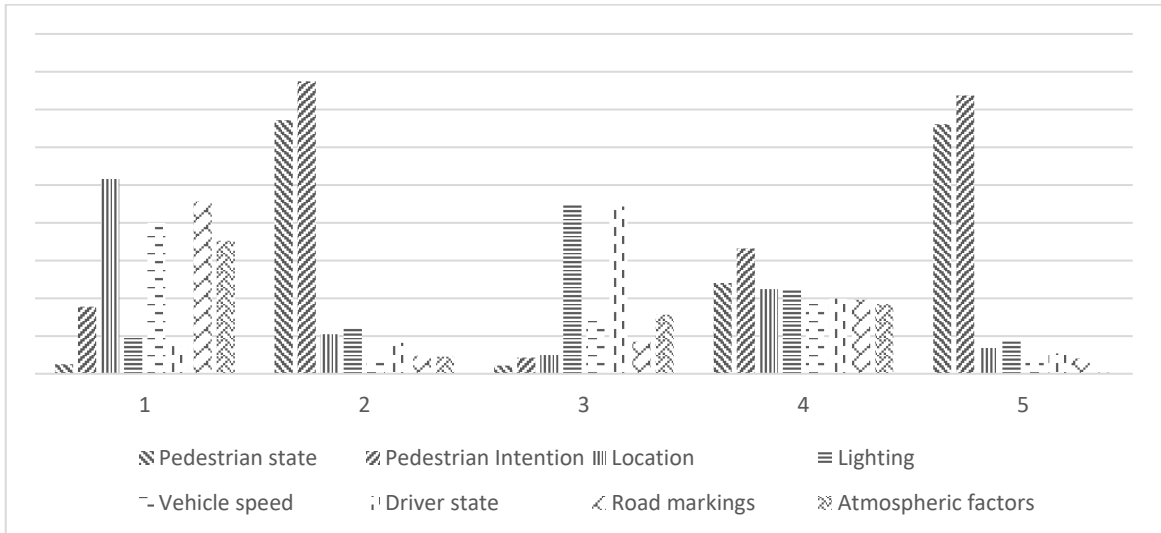


Figure 3. Variables contribution to dimension 2

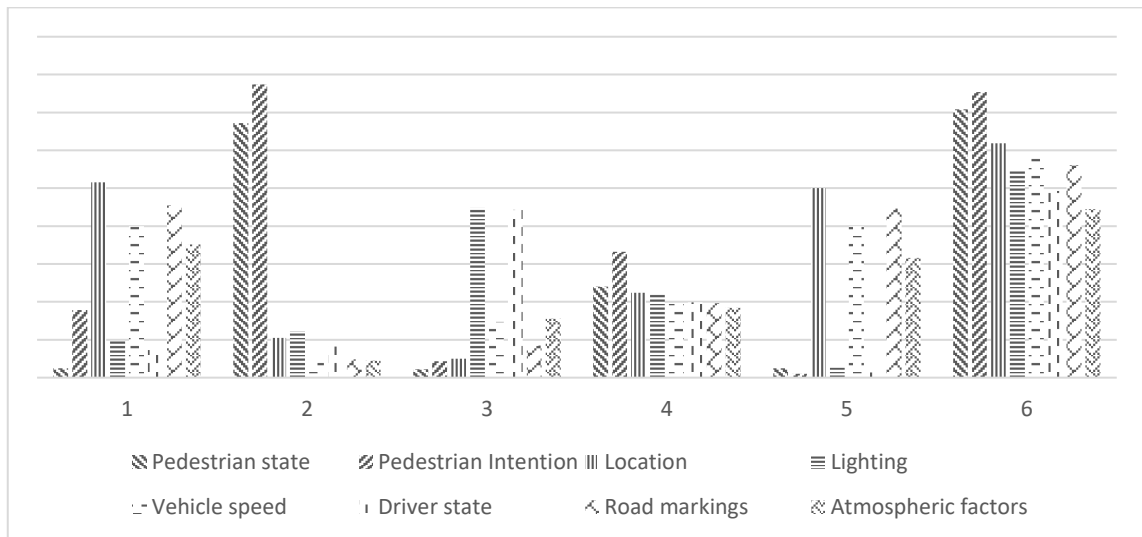


Figure 4. Variables contribution to dimension 3

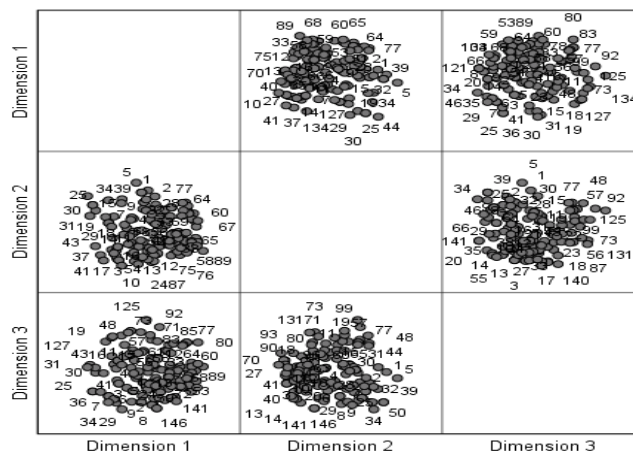


Figure 5. Clustered scatter graph

Component loadings in Table 8 show the coordinates for each item on each dimension, which are plotted in the scatter plot. Table 8 shows how the items related to one another and the three dimensions. The variables with the highest factor loading contribute the highest to the dimension. In this case, a cut-off value of 0.5 was used to select the variables, as indicated in Table 9. Recommends suppression of factor loadings with absolute values less than 0.50 when the sample is not very large. Because have a small sample, defined the cut-off for component loadings as 0.50.

Table 8. Component loadings

| Variables | Dimension | | |
|----------------------|-----------|--------|--------|
| | 1 | 2 | 3 |
| Pedestrian state | | 0.813 | 0.149 |
| Pedestrian intention | | 0.858 | -0.084 |
| Location | -0.707 | -0.262 | -0.224 |
| Lighting | -0.161 | 0.294 | 0.658 |
| Vehicle speed | 0.632 | 0.203 | -0.383 |
| Driver state | 0.119 | -0.231 | 0.652 |
| Road markings | 0.668 | -0.203 | 0.272 |
| Atmospheric factors | 0.562 | 0.059 | 0.355 |

Table 9. Component loadings with values above 0.5

| Variables | Dimension | | |
|----------------------|-----------|-------|-------|
| | 1 | 2 | 3 |
| Pedestrian state | | 0.813 | |
| Pedestrian intention | | 0.858 | |
| Location | -0.707 | | |
| Lighting | | | 0.658 |
| Vehicle speed | 0.632 | | |
| Driver state | | | 0.652 |
| Road markings | 0.668 | | |
| Atmospheric factors | 0.562 | | 0.355 |

5. DISCUSSION

The primary goal of this study was to investigate road accidents while considering factors such as time, location, atmosphere, traffic environment, and injury severity. Road accidents were investigated to determine the factors that influence pedestrian behavior, particularly when crossing the street. It was then discovered how these identified elements influence pedestrian behavior, specifically pedestrian safety.

Pedestrian surroundings, such as lighting conditions, atmosphere, and weather conditions, influence pedestrians' behavior in many ways [22]. This study indicated that most accidents occur in the daylight (40%), with most of them causing severe injuries (19%). On the other hand, the present study did not represent any significant ratio of accidents in adverse weather conditions. However, it is found that most accidents occur in good weather conditions (49%). These unexpected results could be because of the bad weather, which makes people more careful and slows them down as they cross the street, or because of the slippery roads caused by rain, snow, or fog, which also slows down pedestrians and drivers.

Similarly, poor lighting or dark conditions reduce pedestrians' visual functions, enabling them to make riskier decisions [23]. A study indicated by Jägerbrand and Sjöbergh [24] reported similar results and indicated that light conditions did not affect the vehicle's speed. So, it was suggested that the accidents in the poor lighting conditions do not occur due to pedestrians. But the main reason might be that the driver didn't know how to change the speed of the car.

The condition of pedestrians is an important factor in determining pedestrian accidents. It is found that walking pedestrian encounters most accidents, which accounts for 29% of the total accidents. Moreover, it is estimated that 20% of accidents occur with pedestrians running while crossing the road. These results comply with the findings that pedestrian speed can change the visual perception of moving objects. A study conducted by Oudejans *et al.* [25] indicated that pedestrians could understand the intent of moving objects while walking. So, they can accurately guess how far away and how fast the vehicles are, which makes them less careful and more vulnerable.

Apart from walking speed, pedestrians' intentions are another distracting state of pedestrians that could result in road crashes. Our study found that the inattentive behavior of pedestrians caused the highest number of accidents (25%). The distracted behavior of the pedestrians means they were either using phones or busy in some other activity. However, the second-highest number of accidents (22%) occurs when pedestrians dart out of somewhere suddenly. According to the Hyman *et al.* [26], the distraction caused by

electronic devices can make pedestrians less attentive. Consequently, they could either change their walking direction or walking speed unintentionally.

The national highway traffic safety administration (NHTSA) [27] stated that drivers' unintentional and distracting behavior is significantly behind different crashes. Our study confirmed the research conducted by the NHTSA and indicated that 27% of accidents occur due to the sudden application of the brake. It might be attributed to the car's driver being busy with other activities, such as talking on the phone. Distracted drivers often take the braking decision at a shorter distance from the zebra crossing and are mostly unable to control the speed variations [28], [29]. But this ratio is expected to go up when cars can drive themselves with little or no help from their drivers.

Location and traffic volume are also found to influence road safety. Traffic volume can affect both pedestrian's [30] and driver's behavior [31] both. Our study indicated that the highest number of accidents (33%) occurred on state highways. Also, when comparing the number of accidents on roads with different amounts of traffic, it was found that 26% of accidents (the most accidents) happened on roads with between 400 and 1,000 vehicles per day. Generally, high-traffic roads restrict pedestrians' chances of crossing the road, reducing their chances of getting into accidents. On the other hand, higher traffic flow in urban areas increases the risk of pedestrian accidents.

6. CONCLUSION

This study aims to identify the factors that lead to road accidents to establish a safe relationship between pedestrians and vehicles. For this purpose, one hundred and fifty road accident data sets are collected. Eleven variables are considered the leading causes of road accidents. To examine the variables, substantially the component loading with a cut-off value of 0.5 is selected in each dimension. Location, road markings, vehicle speed, and atmospheric factors were significantly involved in the first dimension. The pedestrian state and pedestrian intention have a higher percentage in the second dimension, while lighting and driver state are involved in the third dimension. The results of this analysis and the discussion about them can help automakers and policymakers make systems and algorithms that can handle different situations.




REFERENCES

- [1] F. Hafeez, U. O. Sheikh, N. Alkhalidi, H. Z. Al Gami, Z. A. Arfeen, and S. A. Khalid, "Insights and strategies for an autonomous vehicle with a sensor fusion innovation: A fictional outlook," *IEEE Access*, vol. 8, pp. 135162–135175, 2020, doi: 10.1109/ACCESS.2020.3010940.
- [2] D. Mukherjee and S. Mitra, "A comprehensive study on identification of risk factors for fatal pedestrian crashes at urban intersections in a developing country," *Asian Transport Studies*, vol. 6, pp. 1–14, 2020, doi: 10.1016/j.eastjs.2020.100003.
- [3] X. Zhai, H. Huang, N. N. Sze, Z. Song, and K. K. Hon, "Diagnostic analysis of the effects of weather condition on pedestrian crash severity," *Accident Analysis & Prevention*, vol. 122, pp. 318–324, Jan. 2019, doi: 10.1016/j.aap.2018.10.017.
- [4] S. Zang, M. Ding, D. Smith, P. Tyler, T. Rakotoarivelo, and M. A. Kaafar, "The impact of adverse weather conditions on autonomous vehicles: How rain, snow, fog, and hail affect the performance of a self-driving car," *IEEE Vehicular Technology Magazine*, vol. 14, no. 2, pp. 103–111, Jun. 2019, doi: 10.1109/MVT.2019.2892497.
- [5] C. Holland and R. Hill, "The effect of age, gender and driver status on pedestrians' intentions to cross the road in risky situations," *Accident Analysis & Prevention*, vol. 39, no. 2, pp. 224–237, Mar. 2007, doi: 10.1016/j.aap.2006.07.003.
- [6] A. Al Mamun, P. P. Em, M. J. Hossen, A. Tahabilder, and B. Jahan, "Efficient lane marking detection using deep learning technique with differential and cross-entropy loss," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 4, pp. 4206–4216, Aug. 2022, doi: 10.11591/ijece.v12i4.pp4206-4216.
- [7] F. Malin, I. Norros, and S. Innamaa, "Accident risk of road and weather conditions on different road types," *Accident Analysis & Prevention*, vol. 122, pp. 181–188, Jan. 2019, doi: 10.1016/j.aap.2018.10.014.
- [8] H. Elmannai, M. S. Naceur, M. A. Lohmari, and A. AlGami, "A new feature extraction approach based on non linear source separation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 5, pp. 4082–4094, Oct. 2021, doi: 10.11591/IJECE.V11I5.PP4082-4094.
- [9] K. L. Khine and T. S. Nyunt, "Predictive geospatial analytics using principal component regression," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 3, pp. 2651–2658, Jun. 2020, doi: 10.11591/ijece.v10i3.pp2651-2658.
- [10] F. M. Favarò, N. Nader, S. O. Eurich, M. Tripp, and N. Varadaraju, "Examining accident reports involving autonomous vehicles in California," *PLOS ONE*, vol. 12, no. 9, pp. 1–20, Sep. 2017, doi: 10.1371/journal.pone.0184952.
- [11] L. M. Hulse, H. Xie, and E. R. Galea, "Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age," *Safety Science*, vol. 102, pp. 1–13, Feb. 2018, doi: 10.1016/j.ssci.2017.10.001.
- [12] A. Moradi *et al.*, "Factors affecting the severity of pedestrian traffic crashes," *Archives of Trauma Research*, vol. 8, no. 2, pp. 46–56, 2019, doi: 10.4103/atr.atr_6_19.
- [13] F. N. C. Leitão *et al.*, "Factors associated with incidence and mortality by road accidents involving motorcyclists and pedestrians: A rapid systematic review," *Journal of Human Growth and Development*, vol. 32, no. 1, pp. 72–82, Jan. 2022, doi: 10.36311/jhgd.v32i1.2614.
- [14] H. S. Han and G. H. Lee, "Analysis of children traffic accidents around school area using GIS and data mining techniques," *ICIC Express Letters, Part B: Applications*, vol. 11, no. 10, pp. 961–967, 2020, doi: 10.24507/icieb.11.10.961.
- [15] G. I. Crabb and L. Crinson, "The impact of street lighting on night-time road casualties," *Transport Research Laboratory*, pp. 1–16, 2008.
- [16] V. Sekar and A. Jawaharlalnehru, "Semantic-based visual emotion recognition in videos-a transfer learning approach," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 4, pp. 3674–3683, Aug. 2022, doi: 10.11591/IJECE.V12I4.PP3674-3683.
- [17] J. Julham, M. Lubis, A. R. Lubis, A.-K. Al-Khowarizmi, and I. Kamil, "Automatic face recording system based on quick response code using multicam," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 1, pp. 327–335, 2022, doi: 10.11591/ijai.v11i1.pp327-335.




- [18] C. K. Owusu, J. K. Eshun, C. K. O. Asare, and A. A. Aikins, "Identification of road traffic accident hotspots in the cape coast metropolis, Southern Ghana using geographic information system (GIS)," *International Journal of Scientific & Engineering Research*, vol. 9, no. 10, p. 2106, Oct. 2018, doi: 10.14299/ijser.2018.10.07.
- [19] D. Richards, "Relationship between speed and risk of fatal injury: Pedestrians and car occupants," 2010.
- [20] T. Mphela, T. Mokoka, and K. Dithole, "Pedestrian motor vehicle accidents and fatalities in botswana-An epidemiological study," *Frontiers in Sustainable Cities*, vol. 3, pp. 1–13, Jun. 2021, doi: 10.3389/frsc.2021.666111.
- [21] International Transport Forum, "Road safety annual report 2019," 2019. Accessed: Jun. 27, 2022. [Online]. Available: <https://www.itf-oecd.org/road-safety-annual-report-2019>
- [22] R. E. Amiri, C. Katrakazas, and C. Antoniou, "Negotiation and decision-making for a pedestrian roadway crossing: A literature review," *Sustainability*, vol. 11, no. 23, pp. 1–24, Nov. 2019, doi: 10.3390/su11236713.
- [23] P.-S. Lin, Z. Wang, and R. Guo, "Impact of connected vehicles and autonomous vehicles on future transportation," in *Bridging the East and West*, May 2016, pp. 46–53, doi: 10.1061/9780784479810.006.
- [24] A. K. Jägerbrand and J. Sjöbergh, "Effects of weather conditions, light conditions, and road lighting on vehicle speed," *SpringerPlus*, vol. 5, no. 1, pp. 1–17, Dec. 2016, doi: 10.1186/s40064-016-2124-6.
- [25] R. R. Oudejans, C. F. Michaels, B. van Dort, and E. J. P. Frissen, "To cross or not to cross: The effect of locomotion on street-crossing behavior," *Ecological Psychology*, vol. 8, no. 3, pp. 259–267, Sep. 1996, doi: 10.1207/s15326969eco0803_4.
- [26] I. E. Hyman, S. M. Boss, B. M. Wise, K. E. McKenzie, and J. M. Caggiano, "Did you see the unicycling clown? Inattention blindness while walking and talking on a cell phone," *Applied Cognitive Psychology*, vol. 24, no. 5, pp. 597–607, Oct. 2009, doi: 10.1002/acp.1638.
- [27] C. A. Kahn, "The economic and societal impact of motor vehicle crashes, 2010 (revised)1," *Annals of Emergency Medicine*, vol. 66, no. 2, pp. 194–196, Aug. 2015, doi: 10.1016/j.annemergmed.2015.06.011.
- [28] N. Baldo, A. Marini, and M. Miani, "Drivers' braking behavior affected by cognitive distractions: An experimental investigation with a virtual car simulator," *Behavioral Sciences*, vol. 10, no. 10, pp. 1–16, Oct. 2020, doi: 10.3390/bs10100150.
- [29] S. Yang, "Driver behavior impact on pedestrians' crossing experience in the conditionally autonomous driving context," M.S. thesis, Dept Media Technol. and Interact. Des., KTH Royal Institute of Technol., Stockholm, Sweden 2017.
- [30] C. Ackermann, M. Beggiano, S. Schubert, and J. F. Krems, "An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles?," *Applied Ergonomics*, vol. 75, pp. 272–282, Feb. 2019, doi: 10.1016/j.apergo.2018.11.002.
- [31] C. Lehsing *et al.*, "Effects of simulated mild vision loss on gaze, driving and interaction behaviors in pedestrian crossing situations," *Accident Analysis & Prevention*, vol. 125, pp. 138–151, Apr. 2019, doi: 10.1016/j.aap.2019.01.026.

BIOGRAPHIES OF AUTHORS






Farrukh Hafeez    received the B.E. degree in Electronics and M.E degree in Industrial Electronics from NED University of Engineering and Technology Karachi, Pakistan in 2005 and 2009 respectively. He is currently pursuing the Ph.D. in electrical engineering at Universiti Teknologi Malaysia (UTM). His research interest include computer vision, machine learning, pedestrian intention prediction, sensor fusion. He can be contacted at email: hafeez@graduate.utm.my.






Usman Ullah Sheikh    received his BEng degree (2003) in electrical and mechatronics engineering, the M.Eng degree (2005) in telecommunications engineering and PhD degree (2009) in image processing and computer vision from Universiti Teknologi Malaysia. His research work is mainly on computer vision and embedded systems design. He is currently a senior lecturer at Universiti Teknologi Malaysia working on image processing for intelligent surveillance systems. He is a member of IET and IEEE and has published works in many journals and conferences. He can be contacted at email: usman@fke.utm.my.






Saud Al-Shammari    received his bachelor degree (2010) and Master degree (2012) in electrical engineering from Purdue University Calumet, USA, and PhD degree (2021) in electrical engineering from the University of Sheffield, UK. His primary research interests are power systems, renewable energy, electrical vehicles, and magnetic gearings. He is currently working as an assistant professor at Jubail industrial College. He can be contacted at email: alshammari_s@rcjy.edu.sa.






Muhammad Hamid    received his BEng. degree (1999) in Electrical Engineering from NED University of Engineering Technology Karachi Pakistan and M.S.degree (2002) also in Electrical Engineering from King Fahd University of Petroleum and Minerals KSA. His research work is mainly in embedded system design, control and automation. He is currently working as Lecturer at Qatar Community College. He has published work in research journals and conferences. He can be contacted at email: sadekmh@gmail.com.



Abdul Baqi Khan Khakwani    received his BSc degree (1987) in Electrical Engineering from University of Engineering and Technology, Lahore, Pakistan and MSc (1994) in Information Systems Engineering from London South Bank University, UK. His Primary research interests are Electrical Power System Analysis, Data Communications and Networks. He is currently working as Lecturer at Jubail Industrial College, Kingdom of Saudi Arabia. He can be contacted at email: baqia@rcjy.edu.sa.



Zeeshan Ahmad Arfeen    is an Assistant Professor at “The Islamia University of Bahawalpur”, Pakistan in the Electrical Engineering Department. He is a doctorate scholar from University Technology Malaysia (UTM) Skudai Johor Bahru, Johor, Malaysia. He is a senior member of the IEEE organization NJ. USA, member of IEEE Power and Energy Society (PES), life member IEEE Pakistan, and member of “The Institution of Engineers Pakistan” (IEP). Further, he is also a life member of Pakistan Engineering Council. He is an author of many high indexed Journals and a potential reviewer of more than 160+ papers ranked in ISI, Scopus, and Conference proceeding Journals. He was awarded Partial doctoral scholarship from Higher Education Commission (HEC) Pakistan. He also got an International doctoral fellowship (IDF) and UTM Young Scholarship Grant from University Technology Malaysia (UTM). His one of the papers is the most downloaded paper in 2019 and 2020 and the top 10 cited paper in year 2020 by Wiley publisher. He can be contacted at email: zeeshan.arfeen@iub.edu.pk.