Gait recognition

P.sri Hari

UG Student at Amity University Madhya Pradesh

ABSTRACT

Human gait identification inadvertently detects an individual based on the pedestrian's walk, whereas biometrics such as fingerprints, retinas, palms, and voice recognition need subject consent and physical attention. The goal of this paper is to synthesize prior studies and related work on the human gait recognition system, as well as to develop algorithms that help recognize pedestrians without requiring permission or assistance from the pedestrian. It takes a video of the subject from the camera, turns it to a frame of a still picture, extracts the silhouette using feature extraction techniques, and produces a database to train the silhouette using principal component analysis. It's possible to make it happen. The picture that is entered is the compared to an image that already exists in the database.

Keywords: Gait recognition , machine learning , CASIA Dataset A

1. INTRODUCTION

Identification is one of the maximum crucial components in security. Biometrics is one of the strategies that may be used to identify an individual. For example, fingerprint popularity is used for figuring out human beings from every different via way of means of the use of their fingerprints. In addition to fingerprint popularity, different biometrics consist of ear, vein, retina and gait popularity. Gait popularity, figuring out a person’s frame motion, is likewise a method of biometrics.

In gait analysis, a person’s motion describes non-public manner of strolling and meaning it may be used for figuring out a person. Gait popularity is a biometric approach this is used for figuring out organic and behavioural specification. Gait popularity generation strategies divide into; first one is holistic-primarily based totally technique and the second is model-primarily based totally technique. Holistic based technique is based on extracting statistical capabilities of motion based at the same time as model-primarily

based totally technique identifies frame elements to create a 3-D gait model. In this paper, it became attempted to make the popularity method via way of means of running with many sports primarily based totally on strolling interest and gait popularity generation became favored as an identity machine. To this end, our goal is to discover device mastering strategies and apprehend their professionals and cons on using gait styles for authenticating users.
This machine is one of the distinctive attributes. Open supply database CASIA Dataset A is used in this project .

2. DATASET

On December 10, 2001, Dataset A (formerly the NLPR Gait Database) was produced, which included 20 people. Everyone gets 12 picture sequences, four for each of the three image plane directions (parallel, 45 degrees, and 90 degrees). The duration of each sequence varies depending on the walker's speed, but it must fall between 37 and 127. Dataset A is around 25 megabytes in size, with 3000 photographs in the database.

3. RESEARCH METHODOLOGY

This paper intends to employ different classification techniques which are Linear Regression, Logistic Regression, Random Forest, Support Vector Machine, Decision Tree, Lasso and Ridge Classification models and some ensemble techniques with the gait recognition dataset.

4. Algorithm Used

a) Logistic Regression:

Logistic regression can only be applied if the target class has categorical values. As the aim was to identify the person from there walking style

```
[ ] #split data
    from sklearn.model_selection import train_test_split
    x_train,x_test,y_train,y_test = train_test_split(flat_data,target,test_size=0

[ ] from sklearn.linear_model import LinearRegression
    from sklearn.metrics import mean_squared_error, r2_score

[ ] linear = LinearRegression()

    print(x_train)

[ ] print(y_train)

    [0 0 2 ... 2 0 1]

[ ] print(y_test)

[ ] print(x_test)

[ ] linear.fit(x_train, y_train)

    LinearRegression()

[ ] OLS_predict = linear.predict(x_test)

[ ] linear.score(x_test,y_test)

    0.9204622930495003
```

b) Decision Tree:

Decision tree is one of the most commonly used machine learning classifiers. Taking the best suitable attribute at the root, this algorithm breaks the dataset into partitions. The goal of the partition is to unmix the dataset. The splitting iterates until eventually the partitions group the data such that they are homogeneous.

```
[ ] #Fitting Logistic Regression to the training set
    from sklearn.svm import SVC

[ ]

[ ] svclassifier = SVC(kernel = 'linear')
    svclassifier.fit(x_train, y_train)

    SVC(kernel='linear')

[ ] #Predicting the test set result
    y_pred_S = svclassifier.predict(x_test)
    y_pred_S
    r2_score(y_pred_S, y_test)

    0.9722776900315356
```

c) Random Forest (RF):

Random forest creating a number of decision trees at training and give the output based on the most votes become. This technique used for classification and regression tasks.

```
[ ] #Fitting Random Forest classifiers to the training set
    from sklearn.ensemble import RandomForestClassifier
    classifier_rf = RandomForestClassifier(n_estimators=200, criterion="e
    classifier_rf.fit(x_train, y_train)
    y_pred_rf = classifier_rf.predict(x_test)

[ ] #Creating the Confusion matrix
    from sklearn.metrics import confusion_matrix
    cm_rf = confusion_matrix(y_test, y_pred_rf)
    r2_score(y_pred_rf, y_test)

    0.9646777743724053
```

d) Lasso Classifier:

LASSO is an acronym for Least Absolute Selection and Shrinkage Operator. The LASSO imposes a constraint on the sum of the absolute values of the model parameters, where the sum has a specified constant as an upper bound. This constraint causes regression coefficients for some variables to shrink towards zero.

```
[ ] from sklearn.linear_model import Lasso
    model_lasso = Lasso(alpha= 0.001)
    model_lasso.fit(x_train, y_train)

    Lasso(alpha=0.001)

[ ] pred_lasso = model_lasso.predict(x_test)
    print(np.sqrt(mean_squared_error(pred_lasso, y_test)))

    0.2421188281285823

[ ] r2_score(y_test,pred_lasso)

    0.9117352439257179
```

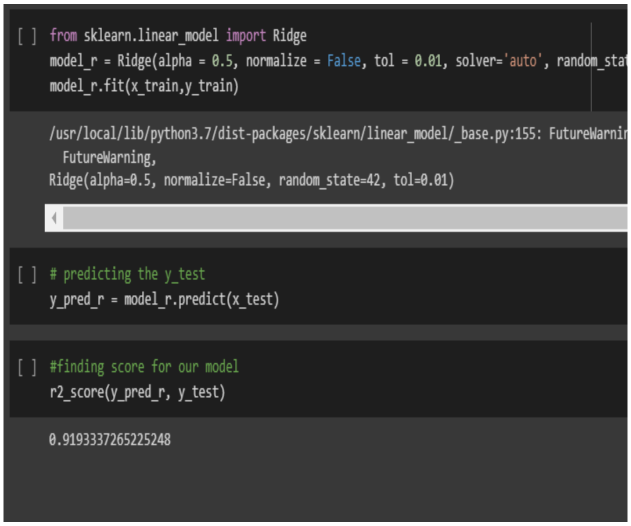
e) Linear Regression

Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on the kind of relationship between dependent and independent variables they are considering, and the number of independent variables getting used.

```
[ ] #split data
    from sklearn.model_selection import train_test_split
    x_train,x_test,y_train,y_test = train_test_split(flat_data,target,test_size=0.3)

[ ] from sklearn.linear_model import LinearRegression
    from sklearn.metrics import mean_squared_error, r2_score

[ ] linear = LinearRegression()

    print(x_train)

[ ] print(y_train)

    [0 0 2 ... 2 0 1]

[ ] print(y_test)

[ ] print(x_test)

[ ] linear.fit(x_train, y_train)

    LinearRegression()

[ ] OLS_predict = linear.predict(x_test)

[ ] linear.score(x_test,y_test)

    0.9204622930495003
```

f) Ridge Classifier:

The Ridge Classifier, based on Ridge regression method, converts the label data into [-1, 1] and solves the problem with regression method. The highest value in prediction is accepted as a target class and for multiclass data multi-output regression is applied.



5. RESULTS

Models	Accuracy
Linear Regression	92%
Logistic Regression	97%
Random Forest	96%
Decision Tree	97%
Lasso Classifier	91%
Ridge Classifier	92%

6. CONCLUSION AND FUTURE WORK

Human gait is a distinctive feature of an individual that is decided by his/her weight, limb length, footwear, and posture in addition with his/her distinctive motion. Gait analysis is widely used as a biometric measure to identify either a person or a subject.

Dataset A (formerly the NLPR Gait Database) was produced, which included 20 people. Everyone gets 12 picture sequences, four for each of the three image plane directions (parallel, 45 degrees, and 90 degrees). The duration of each sequence varies depending on the walker's speed, but it must fall between 37 and 127. Dataset A is around 25 megabytes in size, with 3000 photographs in the database. Our total accuracy with Random Forest classifier is above 96%.

The future work aims to create a better system by using our own dataset to achieve higher accuracy and improve the speed of implementation. We believe that such a system as the one studied in this work could be very useful in real life scenarios for recognizing suspicious people or missing people in crowded areas such as shopping malls, airports, concert venues or cinemas.

7. REFERENCES

[1] O. Sharma, S.K. Bansal, Gait recognition system for human identification Int. J. Innov. Technol. Explor. Eng. 3 (2013) 217–220.

[2] C. Zhang, W. Liu, H. Ma, H. Fu, Siamese neural network-based gait recognition for human identification, in: 2016 IEEE Int. Conf. Acoust. Speech Signal Process., Shanghai, China, 2016, pp. 2832–2836, https://doi.org/10.1109/ICASSP.2016.7472194.

[3] I. Gaba, S.P. Ahuja, Gait analysis for identification in: 2014 IEEE Int. Conf. MOOC, Innov. Technol. Educ., 2014, pp. 122–127, https://doi.org/10.1109/MITE.2014.7020254 .

[4] Y. Liang, C.-T. Li, Y. Guan, Y. Hu, Gait recognition based on the golden ratio, EURASIP J. Image Video Process. 22 (2016), https://doi.org/10.1186/s13640- 016-0126-

[5] Z. Wu, Y. Huang, L. Wang, X. Wang, T. Tan, A comprehensive study on crossview gait based human identification with deep CNNS, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2016) 206–226.

[6] F.M. Castro, M.J. Marín-Jiménez, N. Guil, N.P. De La Blanca, International workconference on artificial neural

[7] B.N. Narayanan, O. Djaneye-Boundjou, T.M. Kebede, Performance analysis of machine learning and pattern recognition algorithms for malware classification, in: 2016a IEEE Natl. Aerosp. Electron. Conf. Ohio Innov. Summit, 2016, pp. 338–342, https://doi.org/10.1109/NAECON.2016. 7856826.

[8] B.N. Narayanan, R.C. Hardie, T.M. Kebede, Analysis of various classification techniques for computer aided detection system of pulmonary nodules in CT, in: 2016b IEEE Natl. Aerosp. Electron. Conf. Ohio Innov. Summit, 2016, pp. 88– 93, https://doi.org/10.1109/NAECON.2016.7856779.

[9] B.N. Narayanan, R.C. Hardie, T.M. Kebede, Performance analysis of a computer-aided detection system for lung nodules in CT at different slice thicknesses, J. Med. Imaging 5 (2018) 14504, https://doi.org/10.1117/1. JMI.5.1.014504.

[10] CASIA DATASET A [http://www.cbsr.ia.ac.cn/users/szheng/?page_id=71]