



TECHNISCHE UNIVERSITÄT
BERGAKADEMIE FREIBERG

Die Ressourcenuniversität. Seit 1765.

Fakultät für Maschinenbau, Verfahrens- und Energietechnik
Institut für Automatisierungstechnik

Seminararbeit

Moderne Methoden der Bildverarbeitung am Beispiel Gesichtserkennung

Anwendung von Informations- und Automatisierungssystemen

Samuel Dressel und Hannes Krumbiegel

25. Oktober 2019

Betreuer/1. Korrektor:
Prof. Andreas Rehkopf

Inhaltsverzeichnis

1	Einleitung	3
2	Grundlagen	3
3	Gesichtserkennung (Face Detection)	4
3.1	Genetische Algorithmen	4
3.2	Neuronale Netzwerke	4
3.3	Viola-Jones-Algorithmus	4
3.3.1	Vorverarbeitung	4
3.3.2	Haar-like-Features	4
3.3.3	Integralbilder	5
3.3.4	Feature-Auswahl	6
3.3.5	Kaskade	6
3.3.6	Nachverarbeitung	6
3.3.7	Verwendung des Algorithmus unter Nutzung von OpenCV	6
4	Gesichtsidentifizierung (Face Recognition)	7
4.1	Methoden der Gesichtserkennung	7
4.1.1	Hauptkomponentenanalyse (PCA) mit Eigengesichtern	7
4.2	Anwendung	8
4.3	Kritik und Probleme	9
5	Zusammenfassung und Ausblick	9
	Literatur	10

1 Einleitung

2 Grundlagen

Einführend zum besseren Verständnis des Themenkomplexes ist es wichtig, einige wesentliche Begriffe zu definieren und außerdem voneinander abzugrenzen. Zunächst gilt es, zwischen der Lokalisation eines Gesichtes an sich und der Zuordnung des Gesichts zu einer bestimmten Person zu unterscheiden. Im ersten Fall wird geprüft, ob und wo ein Gesicht zu sehen ist, im zweiten, um wen es sich handelt [1]. Wir sprechen somit zum einen von der Gesichtserkennung und zum anderen von der Gesichtsidentifizierung. Im englischen Sprachgebrauch unterteilt man den Teil der Gesichtsidentifizierung zusätzlich noch in eine menschliche Erkennung (*face perception*) und in eine maschinelle Erkennung (*face recognition*). In dieser Arbeit wird sich jedoch lediglich auf die maschinelle Gesichtsidentifizierung beschränkt.

3 Gesichtserkennung (Face Detection)

Verschiedene Methoden können verwendet werden, um zu entscheiden, ob ein Bild ein Gesicht enthält. Die Möglichkeiten beinhalten u. a. neuronale Netze, genetische Algorithmen oder speziell für das Problem entwickelte Lösungsansätze. Die Gesichtserkennung ist ein rechnerisch einfacheres Problem als die in Kapitel 4 behandelte Gesichtsidentifizierung, allerdings besteht beispielsweise bei Fotokameras der Wunsch, auf wenig leistungsfähigen Prozessoren die Aufgabe in Echtzeit durchzuführen, um einen schnellen Fokus auf die Gesichter zu ermöglichen. Dieser Anwendungsfall erfordert einen möglichst effizienten Algorithmus.

Der folgende Abschnitt soll einen kurzen Abriss über die verschiedenen Möglichkeiten der Gesichtserkennung geben und exemplarisch einen besonders effizienten Vertreter, den 2001 von Viola und Jones veröffentlichten Viola-Jones-Algorithmus, im Detail vorstellen [2].

3.1 Genetische Algorithmen

3.2 Neuronale Netzwerke

3.3 Viola-Jones-Algorithmus

Der Viola-Jones-Algorithmus ist ein häufig zur Gesichtserkennung eingesetzter Algorithmus. Er verarbeitet in Graustufen umgewandelte Bildausschnitte und nutzt sogenannte *Haar-like-Features*, schwarz-weiße Rechtekanordnungen, um in den Ausschnitten Gesichter zu identifizieren. Diesen Prozess optimiert er mithilfe von verschiedenen Techniken, z. B. den des Integralbilds und einer Attention-Kaskade. Das Resultat ist ein effizienter Algorithmus, der sich auch von schwachen Rechnern in Echtzeit berechnen lässt.

3.3.1 Vorverarbeitung

Der Viola-Jones-Algorithmus operiert ausschließlich auf der Grundlage von Graustufenbildern, weshalb der er im ersten Schritt Farbbilder in diese umwandelt.

Der zweite Schritt dient dazu, die Geschwindigkeit des Algorithmus zu erhöhen. Ein Bild, das weniger Pixel enthält, kann schneller verarbeitet werden, weshalb das Herunterskalieren des Bildes die Performance erhöht. Sehr kleine Gesichter können durch die Reduktion der Auflösung verloren gehen, aber in Anwendungsfällen, wo dies akzeptabel ist, ist der Schritt sinnvoll.

Der dritte Schritt besteht darin, iterativ alle möglichen Bildausschnitte auszuwählen, die ein Gesicht enthalten könnten und demzufolge von einem Gesichtserkennungs-Unteralgorithmus darauf überprüft werden sollen. Es ist beispielsweise möglich, dass das gegebene Bild eine einzige Großaufnahme eines Gesichts ist, der gesuchte Bildausschnitt betrifft hier das ganze Bild. In anderen Fällen enthält ein Bild viele kleine Gesichter: Um diese zu finden, muss der Algorithmus alle Bildausschnitte, die so groß sind wie die gesuchten Gesichter, überprüfen. Letztendlich überprüft der Algorithmus alle Bildausschnitte vorher festgelegter Größen, die von sehr groß bis klein variieren.

3.3.2 Haar-like-Features

Ein Haar-like-Feature besteht aus mehreren schwarzen oder weißen Rechtecken und hat das Ziel, markante dunkle bzw. helle Teile eines Gesichts abstrakt abzubilden. Der Wert eines solchen Features beschreibt, wie gut es auf einen Teil eines Graustufenbildes passt.

Ein Feature, das aus einem schwarzen Rechteck über einem weißen Rechteck besteht, wird einen hohen Wert haben, wenn das schwarze Rechteck über die Augenpartie und das weiße

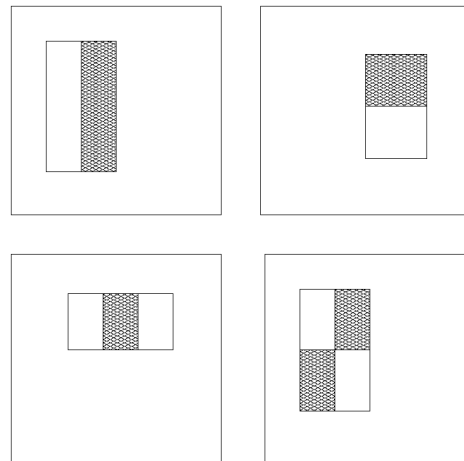


Abb. 1: Mehrere mögliche Haar-like-Features

über Nase und Wangen gelegt wird (siehe Abbildung 2). Bei Teilen des Bildes, die kein Gesicht enthalten, ist dies selten der Fall. Bei einem Bildausschnitt, bei dem viele - für die Gesichtserkennung ausgewählte - Features hohe Werte aufweisen, handelt es sich wahrscheinlich um ein Gesicht.

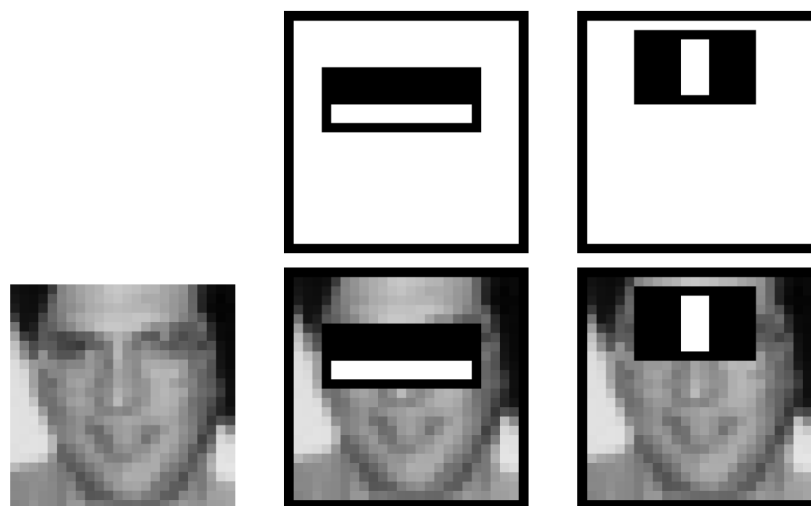


Abb. 2: Zwei Haar-like-Features, die ein Gesicht sehr gut beschreiben.

Der Wert eines solchen Features berechnet sich aus der Summe der Helligkeitswerte der Pixel unter dem schwarzen Rechteck(en), von denen die Summe der Helligkeitswerte unter den weißen Rechteck(en) abgezogen wird. Ein weißes Pixel hat einen Helligkeitswert von 0, ein schwarzes Pixel entspricht dem Wert 255. Ein Feature erreicht dementsprechend seinen Maximalwert, wenn unter seinen weißen Rechtecken nur weiße Pixel und unter seinen schwarzen Rechtecken nur schwarze Pixel liegen.

3.3.3 Integralbilder

Während des Viola-Jones-Algorithmus müssen die Werte von sehr vielen Haar-like-Features berechnet werden. Eine einfache Implementierung für die Berechnung des Werts eines weißen oder schwarzen Rechtecks bestünde aus einer doppelten For-Schleife, die zwischen den Grenzen

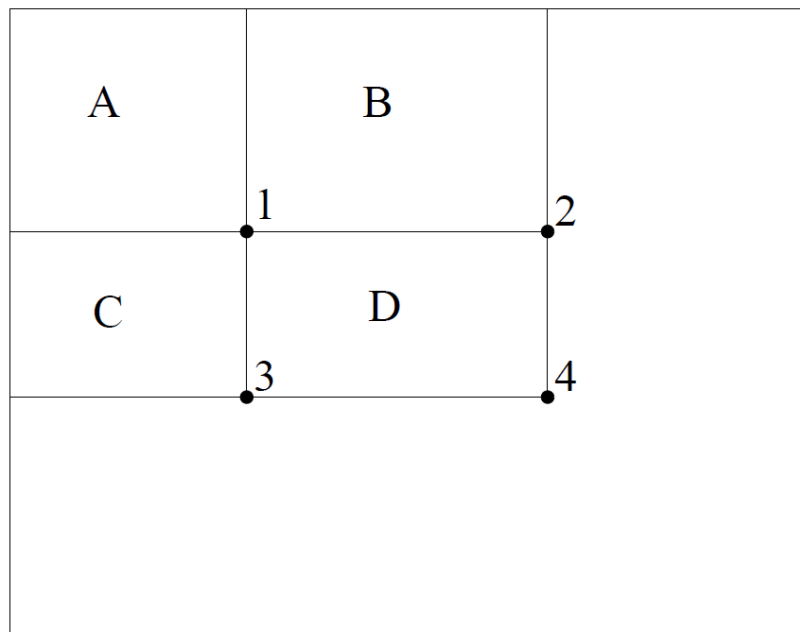


Abb. 3: Eine Illustration eines Integralbildes

der x- bzw. y-Achse iteriert und alle Pixelwerte aufaddiert. Ein Rechteck der Dimension 10x10 würde somit 100 Additionen erfordern.

Viola und Jones entwickelten in [2] das Konzept des Integralbilds, das es ermöglicht, die Anzahl der Rechenoperationen stark zu reduzieren. Das Integralbild ist eine Matrix, die an der Position jedes Pixels des Bildes nicht den Wert des Pixels, sondern die Summe aller Pixelwerte speichert, die links oberhalb des Pixels liegen. Abbildung 3 verdeutlicht dies: Der Eintrag an Position 1 speichert z. B. die Summe der Pixel der Fläche A.

Das Integralbild ermöglicht es, die Summe der Pixel in einem beliebigen Rechteck nur durch 4 Rechenoperationen zu ermitteln.

3.3.4 Feature-Auswahl

3.3.5 Kaskade

3.3.6 Nachverarbeitung

3.3.7 Verwendung des Algorithmus unter Nutzung von OpenCV

4 Gesichtsidentifizierung (Face Recognition)

Das Gebiet der Gesichtsidentifizierung umfasst alle Methoden und Techniken, mit der es möglich ist, basierend auf visuellen Material Menschen eindeutig zu identifizieren. Die Gesichtsidentifizierung baut dabei auf der Gesichtserkennung auf. Insbesondere in den letzten Jahren wurden nach und nach zahlreichen Algorithmen und Methodiken entwickelt, um Gesichter maschinell zu identifizieren. Insbesondere im Hinblick auf die Automatisierung der Gesichtsidentifizierung stellt erst das Jahr 1973 den Startpunkt dar, als Takeo Kanade in seiner Dissertation das erste automatisierte System zur Gesichtsidentifizierung vorstellte [3]. Jedoch konnte gerade aufgrund der begrenzten Rechenleistung zu dieser Zeit dieses System noch längst nicht effizient arbeiten. Ein wesentlicher Fortschritt wurde erst im Jahr 1990 durch die Forscher Kirby und Sirovich erzielt, welche die Hauptkomponentenanalyse (Karhunen-Loève-Transformation) verwendeten [4]. Hierbei wird eine Vielzahl von generierten Variablen auf eine geringe Zahl aussagekräftiger Parameter reduziert. Dadurch wird gleichzeitig auch die Komplexität und der Rechenaufwand verringert. Weiterhin stellte die Arbeit von Matthew Turk und Alex Pentland einen weiteren Meilenstein dar, welche sogenannte Eigengesichter (Eigenvektoren) zur Gesichtsidentifizierung nutzten [5]. Diese Techniken werden auch noch in den derzeitigen State-of-the-Art-Methoden genutzt.

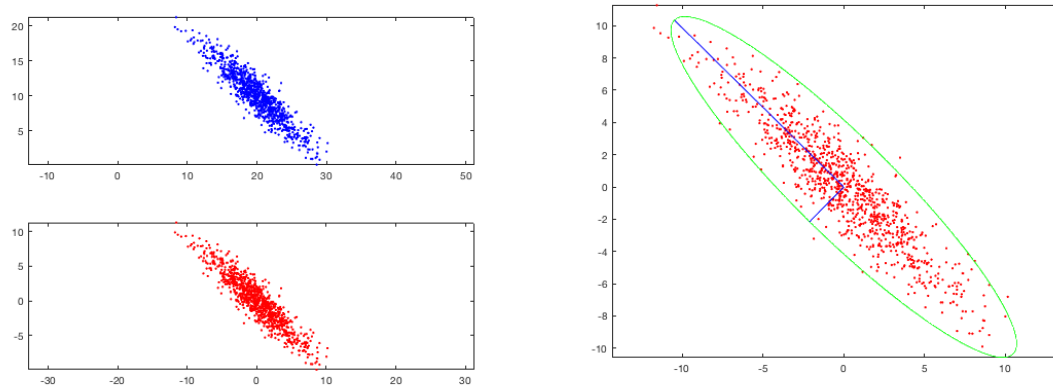
Im Folgenden soll auf eine Reihe dieser Methoden und Herangehensweisen eingegangen, einige Anwendungsbeispiele genannt und die Problematiken der Gesichtsidentifizierung thematisiert werden.

4.1 Methoden der Gesichtserkennung

Grundlegend können alle Methoden der Gesichtsidentifizierung in zwei wesentliche Schritte eingeteilt werden. Der erste Schritt ist die Extraktion und Selektion von Merkmalen aus dem visuellen Input; der zweite Schritt stellt die Klassifizierung der Merkmale dar [6]. Desweiteren können Techniken zur Gesichtsidentifizierung in die Art der verwendeten Merkmale unterteilt werden. Diese können entweder geometrische Features sein (Form der Augen, Position der Nase, ...) oder fotometrische Features sein (Texturen, Farbwerte, ...).

4.1.1 Hauptkomponentenanalyse (PCA) mit Eigengesichtern

Das erste Verfahren, welches im Rahmen dieser Ausarbeitung näher betrachtet werden soll, ist die Hauptkomponentenanalyse (Principal Component Analysis, PCA) mit Eigengesichtern. Die Idee der Gesichtsidentifizierung mit PCA ist es, vorhandenes Bildmaterial so zu reduzieren, dass es sich für eine Auswertung eignet, ohne dabei die wesentlichen Merkmalsinformationen zu verlieren [7]. Man spricht hier von einer Dimensionsreduzierung. Die Hauptkomponentenanalyse an sich beinhaltet noch nicht den kompletten Prozess der Gesichtsidentifizierung, sondern stellt vielmehr ein Tool dar, welches von finalen Methoden im ersten Schritt der Selektion genutzt wird. Dies ist vor allem im Hinblick auf die Effizienz dieses Prozesses von großer Wichtigkeit. Um die Funktionsweise der Hauptkomponentenanalyse mit Eigengesichtern zu verstehen, kann man sich die einzelnen Merkmale als zweidimensionale Punktwolke vorstellen (Abbildung 4a). Eine Dimensionsreduzierung wird den Datensatz in diesen Fall von einer zweidimensionalen Ebene in eine eindimensionale Ebene transformieren [8]. Wie auch bei allen anderen Algorithmen zur Dimensionsreduzierung geschieht dies im Fall der Hauptkomponentenanalyse durch das Finden einer Hyperebene, auf welche die einzelnen Punkte projiziert werden können. Bei der Hauptkomponentenanalyse wird die Hyperebene



(a) Zweidimensionale Punktwolke - beide Datensätze sind identisch mit dem Unterschied, dass das rot gefärbte Diagramm nullzentriert ist

(b) Durch Bestimmung der Eigengesichter generierte Achsen; die längere der beiden blauen Achsen stellt dabei die Achse mit einer maximalen Streuung der Daten da, wenn die Punktwolke auf diese projiziert wird

Abb. 4: Reduzierung der Dimension einer zweidimensionalen Punktwolke

so gewählt, dass bei einer Projektion alle Punkte eine maximale Streuung erzielen. Es wird sozusagen die Achse der maximalen Varianz gesucht. Im Fall der Punktwolke in Abbildung 4 ist dies die längere blaue Linie. Um diese Linie zu finden, kommen nun die sogenannten Eigengesichter zum Tragen. Diese Eigengesichter sind nichts anderes als die Eigenvektoren der Kovarianzmatrix der gegebenen Datenmenge. Da in dem Fall der Hauptkomponentenanalyse die Achsen der maximalen Varianz generiert werden, werden die wichtigsten Informationen aus der Menge von Merkmalen erhalten. Für den zweiten Schritt der Klassifizierung im Rahmen der Gesichtsidentifizierung wirkt sich eine große Streuung zusätzlich positiv aus.

4.2 Anwendung

Mit den derzeitig vorhandenen Methoden und einer ausreichend hohen Rechenleistung lässt sich die automatisierte Gesichtsidentifizierung vielfältig nutzen und wird auch mehr und mehr in verschiedensten Anwendungsgebieten eingesetzt. Ein erstes großes Anwendungsgebiet ist sicherlich die Nutzung zu Überwachungszwecken und in Verifizierung von Personen in rechtstaatlichem Kontext. Grade in Staaten wie China, Russland oder der USA. Als Beispiel sei hier das „Next Generation Identification System“ des FBI genannt, welches mit einer Genauigkeit von 85 % Personen identifizieren kann [9]. Weiterhin setzt China hier unter Verwendung von Künstlicher Intelligenz neue Maßstäbe im Hinblick auf eine komplette Überwachung [10]. Beispielsweise werden Bürger, die bei Rot über die Ampel gehen, durch Kameras identifiziert und durch Bildschirme öffentlich als Verkehrssünder angeprangert. Hier werden auch die negativen Aspekte des Einsatzes deutlich.

Weiteren Einsatz findet die Gesichtsidentifizierung auch privaten und finanziellen Sektor. Hier reicht die Einsatzreichweite von dem Entsperren von Türen und Smartphones bis hin zum Bezahlen mithilfe der Gesichts-ID. Einen Meilenstein erzielte Facebook mit der Entwicklung von „Deep Face“, welches die Personenidentifizierung mit einer Genauigkeit von 97 % erlaubt [11]. Um diese Genauigkeit zu erzielen, wurde eine neurales Netz mit neun Schichten und 120 Millionen Gewichtungen benutzt und mithilfe von vier Millionen Facebook-Profilbildern trainiert.

4.3 Kritik und Probleme

Die Thematik der Gesichtsidentifizierung im Vergleich zu einer lediglich Gesichtserkennung im Sinne der Unterscheidung von anderen Objekten bringt eine Vielzahl von neuen Problemen mit sich. Und dies längst nicht nur auf der technischen Seite, sondern auch im Hinblick auf ethische Aspekte. Betrachtet man die auftretenden Probleme der technischen Umsetzung, so können unterschiedliche Herausforderungen durch die Verwendung von unterschiedlichen Methoden und Technologien gemeistert werden. Jedoch werden bestimmte Schwierigkeiten noch längst nicht vollständig gelöst. Klassische Methoden, welche ihren Input aus klassischem Bildmaterial beziehen, haben mit einer Reihe ganz unterschiedlicher Probleme zu kämpfen [12]:

- Viele zweidimensionale Methoden liefern nur verwendbare Ergebnisse unter ausreichend guten Lichtverhältnissen. Auch ein großer Kontrast bzw. Variation der Beleuchtung kann die Effizienz dieser Methoden erheblich einschränken.
- Die Verdeckung insbesondere der oberen Gesichtsareale durch eventuelle Gadgets wie Brillen, Hüte oder Kapuzen verringert die Genauigkeit oder kann eine Gesichtsidentifizierung nahezu unmöglich machen. Abhilfe schafft hier die Verwendung von Wärmesignaturen zur Gesichtsidentifizierung.
- Klassische Verfahren, die auf die Textur bzw. die Form des Gesichts zur Identifizierung angewiesen sind, funktionieren nur, wenn ein ausreichend optimaler Blickwinkel auf das Gesicht gegeben ist.
- Bei der Identifizierung von Menschen muss der natürliche Alterungsprozess und damit die Veränderung des Gesichts als ein negativer Faktor beachtet werden. Dies erfordert entweder eine aktuelle Vergleichsdatenbank oder das Verwenden einer hinreichend guten KI, um den Alterungsprozess des Gesichts abbilden zu können.
- Ein weiteres Problem im Hinblick auf eine fotometrische Gesichtserkennung sind Emotionen. Diese können ähnlich wie Verdeckungen durch eine damit einhergehende Gesichtsverzerrung die Identifizierung wesentlich ungenauer machen.

Neben den technologischen Herausforderungen der Gesichtsidentifizierung ergeben sich auch eine Reihe von ethischen Problemen bzw. Kontroversen. Viele Menschenrechtsorganisationen und private Kampagnen zeigen immer wieder Zweifel und äußern Kritik, wenn es um den Einsatz dieser Technologien im Überwachungsbereich geht. Diese Bedenken reichen bis hin zum Gegenstand einer der „totalen Überwachungsgesellschaft“, bei dem Staaten und andere Autoritäten die Fähigkeit haben, die Überwachung ihrer Bürger umfassend auszuweiten. Die Verwendung von Identifizierungstools bietet außerdem Möglichkeiten des Missbrauchs im Hinblick auf die digitale Welt (Social Profiling [13]).

5 Zusammenfassung und Ausblick

Literatur

- [1] Gesichtserkennung. <https://de.wikipedia.org/wiki/Gesichtserkennung>. Zuletzt besucht: 23.10.2019.
- [2] Paul Viola; Michael Jones. Rapid object detection using a boosted cascade of simple features, 2001.
- [3] Takeo Kanade. *Picture processing system by computer complex and recognition of human faces*. PhD thesis, Department of Information Science, Kyoto University, 1973.
- [4] M. Kirby; L. Sirovich. Application of the karhunen-loueve procedure for the characterization of human faces. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1990.
- [5] M. Turk; A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 1991.
- [6] Facial recognition system. https://en.wikipedia.org/wiki/Facial_recognition_system. Zuletzt besucht: 23.10.2019.
- [7] Ana-Maria Sevcenco; Wu-Sheng Lu. Extendend 2-D PCA for Face Recognition: Analysis, Algorithms and Performance Enhancement. In *Face Recognition*. Nova Science Publishers, Inc, 2012.
- [8] Mohit Deshpande. Face Recognition with Eigenfaces. <https://pythonmachinelearning.pro/face-recognition-with-eigenfaces>. Zuletzt besucht: 23.10.2019.
- [9] Next Generation Identification (NGI). <https://www.fbi.gov/services/cjis/fingerprints-and-other-biometrics/ngi>. Zuletzt besucht: 23.10.2019.
- [10] Greg Walton. *China's Golden Shield: Corporations and the Development of Surveillance Technology in the People's Republic of China*. Rights and Democracy, 2001.
- [11] Yaniv Taigman; Ming Yang; Marc Aurelio Ranzato; Lior Wolf. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [12] Asit Kumar Datta; Madhura Datta; Pradipta Kumar Banerjee. *Face Detection and Recognition - Theory and Practice*. CRC Press - Taylor & Francis Group, 2016.
- [13] Muhammad Bilal; Abdullah Gani; Muhammad Ikram Ullah Lali; Mohsen Marjani; Nadia Malik. Social Profiling: A Review, Taxonomy, and Challenges. *Cyberpsychology, Behavior and Social Networking*, Vol. 22, No. 7, 2019.