

Metoder och applikationer för ansiktsigenkänning

Viktor Sjöström

40640

Kandidatavhandling i datateknik

Handledare: Sébastien Lafond

Åbo Akademi

2019

Referat

Datorseende är ett område som fokuserar på att automatisera processer som använder människans vision. Inom datorseende finns det igenkänning, ett av de svårare delområden inom datorseende, som försöker analysera och klassificera objekt på en bild. Specifika egenskaper av objektet kan jämföras med tidigare inlärd information för att identifiera vad objektet representerar. Ett ansikte kan på samma sätt analyseras och identifieras med de unika ansiktsdragen som varje människa har.

Ansiktsigenkänning används dagligen som verifikation och inom säkerhet. Företag som Google, Microsoft och IBM har skapat API:er (programmeringsgränssnitt) för ansiktsigenkänning för kommersiellt bruk. Denna kandidatavhandling behandlar de olika metoderna som används för ansiktsigenkänning. I den andra delen av avhandlingen analyseras olika API:er och verktyg som används inom ansiktsigenkänning.

Innehållsförteckning

| | | |
|--------|---|----|
| 1. | Inledning | 1 |
| 2. | Bakgrund..... | 2 |
| 2.1. | Datorseende | 2 |
| 2.2. | Igenkänning | 2 |
| 2.3. | Biometri..... | 3 |
| 2.4. | Maskinlärning..... | 3 |
| 2.5. | Djupinlärning..... | 4 |
| 2.5.1. | Uppbyggnaden av ett neuralt nätverk | 4 |
| 3. | Ansiktsigenkänning | 6 |
| 3.1. | Ansiktsigenkänningsprocessen..... | 6 |
| 3.2. | Användning | 6 |
| 3.2.1. | Mobila apparater | 7 |
| 3.2.2. | Säkerhet..... | 7 |
| 3.3. | Metoder inom ansiktsigenkänning | 7 |
| 3.3.1. | Egenansikten (Eigenfaces)..... | 8 |
| 3.3.2. | Fisherfaces | 10 |
| 3.3.3. | Viola-Jones..... | 10 |
| 3.3.4. | Faltningsnätverk..... | 12 |
| 3.3.5. | 3D ansiktsigenkänning..... | 13 |
| 4. | Verktyg | 14 |
| 4.1. | MATLAB | 14 |
| 4.2. | OpenCV..... | 14 |
| 4.3. | OpenFace..... | 14 |
| 5. | APIs | 15 |
| 6. | Diskussion..... | 16 |
| 7. | Slutsats..... | 17 |
| 8. | Källor | 18 |

1. Inledning

[beskrivning av ämnet, syfte, sammanfattning av uppläggnig]

Målet med avhandlingen är att ge läsaren en förståelse för hur ett ansiktigenkänningsystem fungerar och skillnaden mellan de olika verktygen för igenkänning och API för ansiktigenkänning.

2. Bakgrund

[I detta kapitel presenteras de olika byggstenarna som är grunden för ansiktsigenkänning.]

2.1. Datorseende

Datorseende, eller Computer Vision på engelska, är ett område där en dator får in data i form av en bild eller en sekvens av bilder och på basis av informationen som datorn identifierat kan utföra en specifik handling. När en människa ser på en bild kan vi nästan omedvetet identifiera alla objekt på bilden och olika egenskaper av varje objekt. En dator ser bilden som en matris av siffror, som ger ett värde för vilken färg som befinner sig i varje pixel av bilden. [1]

Ett problem som uppstår är att bilderna, som är tvådimensionella, är baserade på en tredimensionell värld. Människor kan se på en tvådimensionell bild och förstå vad som bilden föreställer i ett tredimensionellt sammanhang. För att lösa detta problem skapar man en tredimensionell framställning av ett objekt med hjälp av en samling bilder från olika synpunkter.

2.2. Igenkänning

Den svåraste visuella uppgiften inom datorseende är igenkänning och analysen av objekt på en bild. Igenkänning är svårt för datorer eftersom vår värld består av olika objekt med specifika egenskaper och även objekt som tillhör samma klass kan ha stor variation. På grund av den stora variationen av objekten är det viktigt att definiera ett sammanhang för bilden. [1]

Ifall man skulle försöka använda en igenkänningsalgoritm på en bild utan att först upptäcka var objektet man vill känna igen befinner sig, skulle algoritmen med sannolikhet vara långsam och vara benägen att ge fel. Istället används detektorer som hittar regionen där objektet finns före den analyseras djupare av igenkänningsalgoritmer.

2.3. Biometri

Biometri är termen som används för de fysiska särdragen hos människan som är unika för den personen. Eftersom dessa särdrag är unika för varje människa kan de användas som identifiering och igenkänning av personer. Till biometri hör bland annat ögonskanning, ansiktsgigenkänning, röstigenkänning och fingeravtryck. När ett ansikte skall identifieras jämförs bilden med en databas av bilder och ger ett svar på den personen med den högsta sannolikheten.

När biometri används som verifikation, försöker systemet svara på frågan ”Är detta person X?” som användaren påstår sig vara. Systemet jämför sedan användarens biometriska sampel med den mallen som sparats av den ursprungliga användaren och kommer antingen att lyckas eller misslyckas med att hitta ett samband mellan de två bilderna.

2.4. Maskinlärning

Kombinationen med maskininlärning och datorseende har ökat flexibiliteten och noggrannheten av igenkänning av bilder. Maskininlärning är en form av artificiell intelligens, där en dator tar in exempeldata och får en uppfattning av vad det handlar om utan att programmeras direkt. Maskininlärning delas upp i övervakad och oövervakad inlärning. *Oövervakade* algoritmer är då man endast känner till indata och ingen utdata ges för algoritmen. Den *övervakade* maskininlärningsalgoritmen, där maskinen kan förutspå ett utdata baserat på exempelindata, anses vara mer framgångsrik på grund av bättre prestanda och precision jämfört med oövervakad inlärning. Efter maskinen har fått exempel på indata och respektive utdata, kan den själv producera ett utdata utan mänsklig hjälp från ett obekant indata. [2] Det finns dock en risk att man ger för mycket exempelindata till maskinen så att den inte själv kan anpassa till ett indata som är obekant, detta kallas för *överanpassning*.

Exempelindata och utdata paren existerar som en samling av data som kallas för *dataset*. Dessa datasets skapas manuellt av personer som betecknar ett utdata för

varje indata. Det finns datasets som är öppna för allmänheten från forskare som har skapat datasets för objektupptäckning och övriga bildklassificeringsuppgifter.

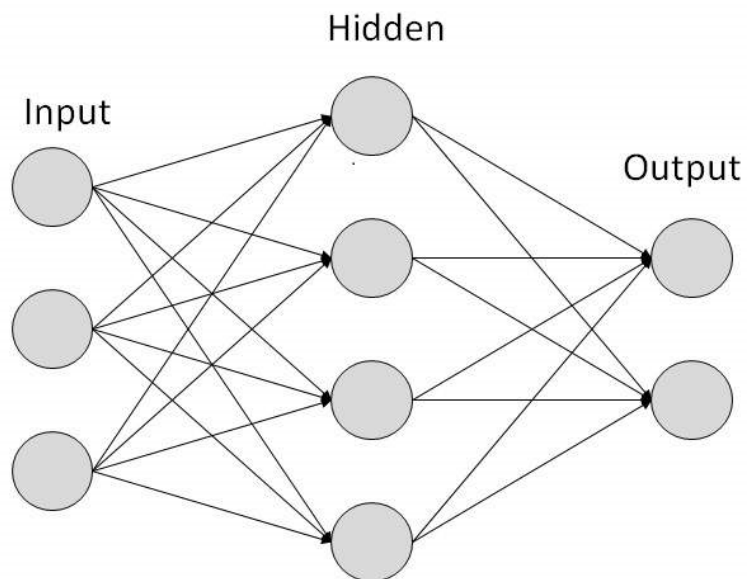
Inom ansiktsigenkänning saknas det dock storskaliga datasets för allmänheten, dessa datasets skapas oftast från bilder på offentliga personer, som exempel IMDB-WIKI dataset som innehåller över 500 000 bilder på ansikten samlat från Imdb och Wikipedia. Googles ansiktsigenkänningsystem FaceNet använder ca. 200 miljoner bilder på 8 miljoner unika identiteter.

2.5. Djupinlärning

Djupinlärning är ett delområde inom maskininlärning, som är ett delområde inom artificiell intelligens. Djupinlärning inspirerades av biologin, en hjärna består av ca. 86 miljarder neuroner som behandlar och överför information i form av elektroniska och kemiska signaler. En hjärna kan kallas för ett biologiskt neuralt nätverk, och inom djupinlärning skapar man ett artificiellt neuralt nätverk som följer huvudegenskaper baserat på hur hjärnan fungerar. Basenheten inom neurala nätverk kallas för en nod, som baserar sig på de biologiska neuronerna, får indata från andra noder och skickar vidare information.

2.5.1. Uppbyggnaden av ett neuralt nätverk

Ett neuralt nätverk är uppbyggt av lager, dessa lager innehåller noder och noderna är kopplade till varandra. Lagerna består av ett indatalager, som placerar indata värden i form av vektorer in i noderna. Sedan finns det ett eller flera gömda lager, där varje neuron i lagret har en koppling med varje neuron från det föregående lagret. Kopplingarna har ett värde för vikt w , som anger styrkan av värdet som neuronerna skickar vidare, vikterna justeras under inlärningsprocessen. Till sist kommer det utdatalager som ger ett eller flera real-värden.



Figur 1. Illustration av ett feedforward neuralt nätverk.

https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_neural_networks.htm

3. Ansiktsigenkänning

3.1. Ansiktsigenkänningsprocessen

Ansiktsigenkänningen sker i allmänhet i fyra steg, men det finns flera varierade metoder. Först måste man använda *ansiktsdetektion*, som segmenterar ansiktet från bakgrunden. Ansiktsdetektionen ger en ungefärlig position av platsen och skalan på ansiktet, men med användningen av *face landmarking* kan man lokalisera ansiktsdrag, som ögonen, näsan och munnen, för att få en mer exakt position av ansiktet.

Sedan utförs en *ansiktsnormalisering* som beskär och transformerar ansiktet från en bild för att minska på störningar som finns på bilden. *Geometrisk normalisering* transformerar ansiktet till en standardstorlek för systemet, genom att beskära bort allt som inte innehåller ett ansikte som hittats under ansiktsdetektionen. En *fotometrisk normalisering* normaliserar ansiktet med tanke på egenskaper som belysning och gråskalan.

Särdragsextraktion används på de normaliserade ansiktena för att extrahera framträdande information som gör att systemet kan skilja på olika ansikten. Dessa extraherade särdrag används sedan för ansiktsmatchning mot en databas med registrerade ansikten. I ett 1:1 matchningssystem så används ett ansikte som indata och den jämförs med ett ansikte som har registrerats som användaren. Detta är ett verifikationssystem och kommer antingen att acceptera eller inte acceptera ansiktet som har gavs. För identifikation används en 1:N matchning, ett ansikte jämförs med N stycken ansikten som befinner sig i en databas. Utdata kommer att skriva ut identiteten av som matchar med högsta förtroendenivån, eller skriva ut okänd person ifall förtroendenivån inte uppnår ett specifikt tröskelvärde. [3]

3.2. Användning

Ansiktsigenkänning används för flera olika ändamål, oftast inom säkerhet, övervakning och verifikation.

3.2.1. Mobila apparater

Användningen av ansiktsgenkänning på smartmobiler har ökat under de senaste åren på grund av den ökande prestandan av apparaterna och den framåtvända kameran som finns på de flesta moderna smartmobilerna. Ansiktsautentisering[4]

3.2.2. Säkerhet

Ansiktsgenkänning används inom säkerhet som ett sätt att identifiera personer och granska ifall systemet känner igen personen. Ansiktsgenkänning som ett system för säkerhet har implementerats på många flygplatser runt om i världen. Igenkänningssystemet jämför indata ansiktena med en databas på misstänkta brottslingar för att alarmera säkerhetspersonalen, som kan manuellt granska situationen vidare. På flygplatser används även ansiktsgenkänning för att granska identiteter utan att behöva en manuell granskning, och personens ansikte jämförs med bilden de har som passfoto.

Till skillnad från andra biometriska metoder så kan ansiktsgenkänning användas på distans och även användas i hemlighet, utan att personen vet att den filmas. Detta har skapat bekymmer, eftersom dessa system kan användas att övervaka stora mängder människor och möjligtvis minska på personlig integritet.

3.3. Metoder inom ansiktsgenkänning

Inom ansiktsgenkänning finns det olika metoder som har utvecklats sen 1973 då den första automatiska ansiktsgenkänningen utfördes. Ansiktsgenkänningsalgoritmerna brukar delas in i tre kategorier, helhetsbaserad, särdragsbaserad och hybrida metoder. I de helhetsbaserade metoderna tas hela ansiktet i beaktande som indata, exempel på dessa metoder är Eigenface, Fisherface och oberoende komponentanalys. Särdragsbaserade metoder fokuserar på ansiktsdrag, som ögonen och näsan, och lokaliserar särdragen. Hybrida metoder använder en kombination av de helhetsbaserade och särdragsbaserade metoderna och används oftast för tredimensionell ansiktsgenkänning. [5]

3.3.1. Egenansikten (Eigenfaces)

Egenansikten baserar sig på idén att en bild x kan komprimeras och rekonstrueras med ett medel-ansikte m , som skapats av ett övnings-set av ansikten.

Med hjälp av Principalkomponentanalys, förkortat PCA, vill man hitta de vektorer inom bilderna som bäst representerar ett ansikte. Dessa vektorer kallas för egenvektorer, men på grund av deras ansiktsliknande utseende, kallas dessa vektorer för *egenansikten*. [6] Utförandet av PCA på en mängd bilder för att få egenansiktena gör enligt följande steg (baserat på [6]):

Först måste man ha ett data set D med en samling av M antal ansikten, som fungerar som övnings set. Bilderna har en storlek på $N * N$, där N är en bestämd mängd pixlar för höjden och bredden.

Varje bild kan representeras som en N^2 vektor eller en kolumn, där varje rad i kolumnen är gråskalevärdet för en pixel i bilden i N^2 rader. Efter detta kan en matris A skapas, med antalet M kolumner för varje ansikte med N^2 pixlar per ansikte.

För att kalkylera egenansikten från matrisen, måste man först skapa ett medelansikte Ψ . Varje övningsansikte representeras av Γ , där $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_K$

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n$$

Efter det justeras ansiktena genom att subtrahera medelansiktet Ψ från varje ansikte Γ .

$$\Phi_i = \Gamma_i - \Psi$$

De justerade ansiktena Φ_M visar de skillnader som finns för varje ansikte jämfört med medeltalsansiktet Ψ .

Kovariansmatrisen C fås av ekvationen $C = AA^T$, där $A = [\Phi_1, \Phi_2, \dots, \Phi_M]$. Matrisen C är av storleken $N^2 * N^2$ vilket bild en krävande operation för typiska bildstorlekar. Ifall antalet bilder M är mindre än storleken N^2 , finns det $M - 1$ meningsfulla egenvektorer istället för N^2 vektorer. Detta kan man göra eftersom de

återstående vektorerna har egenvärdet noll, vilket gör att längden för vektorerna är också noll. Då kan man skapa en $M * M$ matris L , där $L = A^T A$, de normaliserade egenvektorerna u_i och egenvärdena λ_i från C kan då fås genom följande ekvation

$$u_i = \lambda_i^{-1/2} A v_i \text{ för } i = 1, \dots, M$$

Dessa egenvektorer u_i kallas även för egenansikten för data set D . Med hjälp av egenansiktena kan man nu använda de inom ansiktsigenkänning. Området är egenansiktena befinner sig kallas för "face space". Ett indata ansikte Γ transformeras till dess egenansikte komponenter och projiceras in i "face space":

$$\omega_k = u_k^T (\Gamma - \Psi), k = 1, \dots, M$$

Ekvationen kalkylerar vikter ω_k , som bildar en vektor $\Omega^T = [\omega_1, \omega_2 \dots \omega_M]$. Vikterna ger ett värde, högre värden anger en högre liknelse mellan indata ansiktet och egenansiktena. För att svara på frågan "Innehåller denna bild ett ansikte?" kan man räkna ut distansen ϵ mellan den justerade indata ansiktet $\Phi = \Gamma - \Psi$ och projektionen till "face space" $\Phi_f = \sum_{i=1}^M \omega_i u_i$ med:

$$\epsilon^2 = \|\Phi - \Phi_f\|^2$$

Ifall distansen ϵ^2 är mindre än ett tröskelvärde Θ , är indata ansiktet identifierat som ett ansikte. Tröskelvärdet kan justeras beroende på applikationen av algoritmen, högre tröskelvärden har mer precisa resultat men tillåter inte små variationer i ansiktet.

3.3.2. Fisherfaces

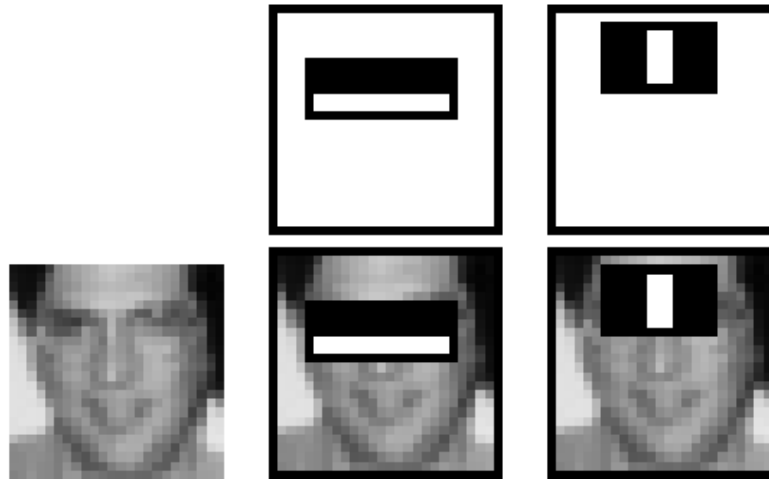
En annan metod för ansiktsigenkänning används med hjälp av Linjär diskriminantanalys, förkortat LDA. Målet med Fisherfaces var att skapa en algoritm för ansiktsigenkänning som skulle känna igen ansikten under varierande belysning och ansiktsuttryck. I artikeln [7] kommer de till sammanfattningen att Fisherfaces har bättre prestanda för bilder som har variationer i belysning och ansiktsuttryck jämfört med PCA, korrelations- och linjära delrumsalgoritmer.

3.3.3. Viola-Jones

Paul Viola och Michael Jones skapade år 2001 ett ramverk för snabb och robust ansiktsupptäckning. Viola-Jones ramverket för att upptäcka ansikten minskade beräkningstiden och behåller hög upptäckningsprecision. Efter ett ansikte har upptäckts, kan man använda igenkänningsmetoder som PCA och LDA för att extrahera särdragen och jämföra resultaten med en databas. Viola-Jones ramverket kan detektera 15 bildrutor per sekund, vilket kan användas för ansiktsupptäckning av en video. [8]

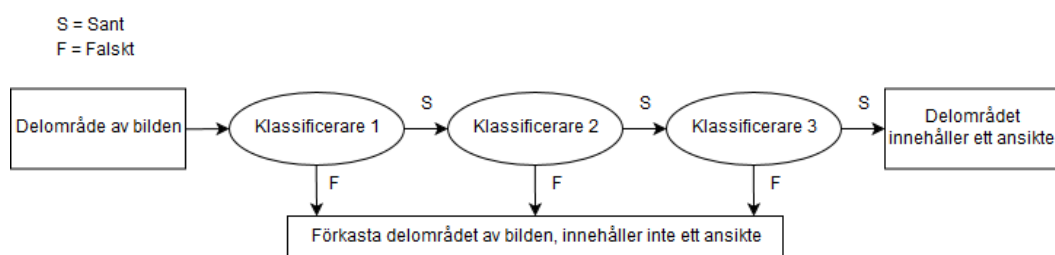
Ramverket baserar sig på konceptet av "boosting" med hjälp av maskininlärningsalgoritm kallad AdaBoost (Adaptive Boosting). Istället för att identifiera alla särdrag från en bild så exkluderas majoriteten av särdragen och endast de kritiska särdragen tas i beaktande. En slutlig klassificerare skapas av en summa av svaga klassificerare som endast fokuserar på ett särdrag. Det existerar över 180 000 rektangulära särdrag för varje delområde på en bild, endast en liten del av särdragen används och den största utmaningen är att hitta dessa särdrag. [8]

AdaBoost algoritmen tränas med exempel indata med bilder på ansikten och bilder som inte innehåller ett ansikte. För varje övningsskede bestäms ett särdrag från 180 000 potentiella särdrag. [8] Figur 2 visar de första och andra särdragen som valts av AdaBoost. Det första särdraget jämför intensitetsskillnaden mellan ögonregionen och övre kinden, det andra särdraget jämför skillnaden mellan ögon och näsryggen.



Figur 2. Olika rektangulära särdrag som filtreras med ett Haar filter <https://medium.com/neuromation-blog/neuronuggets-object-detection-9a7a17f7b2fc>

Särdragen jämför alla delområden av en bild, accepterar områden som klarar av jämförelsen och förkastar delområdet om jämförelsen inte når ett valt tröskelvärde. De första särdragen som jämförs har oftast ett lägre tröskelvärde för accepterande av bilden. Vid detta skede kommer alla ansikten att gå vidare och även områden från bilder utan ansikten, eftersom endast ett särdrag har jämförts. De bilder som inte når tröskelvärdet förkastas direkt och nästa delområde jämförs. Senare särdrag blir mer detaljerade och specifika, vilket leder till att bilder som inte innehåller ett ansikte kommer att förkastas vid något skede. När ett delområde har lyckats med alla särdrags jämförelser, kan man konstatera att det delområdet innehåller ett ansikte (Figur 3).



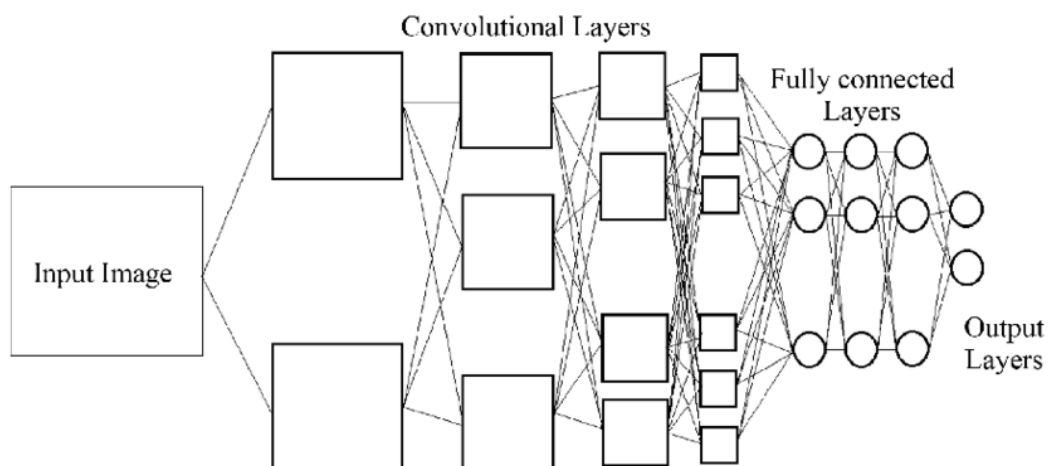
Figur 3. Processen för ansiktsdetektion

Viola-Jones ramverket används ännu idag, fastän många upptäckningssystem använder neurala nätverk. Viola-Jones ramverket är ännu snabb och pålitlig, och kräver inte en kraftig processor vilket gör algoritmen lämplig för mobila apparater, som smarttelefoner och digitala kameror.

3.3.4. Faltningsnätverk

Det finns flera olika arkitekturer för djupa neurala nätverk och Faltningsnätverk, förkortat CNN (eng. Convolutional Neural Networks), används ofta inom igenkänning av visuellt data. I kapitel 1.5.1 beskrevs uppbyggnaden av ett så kallat "Feed-forward multilayer neural network", vilket tar in en en-dimensionerad vektor och transformerar data med ett eller flera gömda lager. Dessa neurala nätverk anpassar sig inte bra till bilder som indata, eftersom informationen av bilden skulle skapa en massiv mängd av kopplingar mellan varje neuron och mängden beräkningar skulle skapa ett ineffektivt igenkänningssystem.

I en CNN kan man strukturera bilderna enligt pixelbredd, pixelhöjd och djup, i form av RGB värden, för att behålla den tredimensionella strukturen av bilden. Arkitekturen för ett CNN kan delas upp i tre grupper: indatalagret, särdragsextraheringslager och klassificeringslager. Indatalagret tar in tredimensionella input i form av höjden, bredden och djupet, vilket oftast representeras som tre lager för RGB kanalerna. Särdragsextraheringslagren består av en sekvens av repeterande lager, faltningslager, ReLU (eng. Rectified Linear Unit) lager och poollager. Dessa lager hittar särdrag i bilderna och skapar mer detaljerade särdrag då det finns flera lager. Klassificeringslagret består av ett eller flera fullständigt anslutna lager, som är anslutna till det tidigare lagret. [8]



Figur 4 Arkitekturen för ett faltningsnätverk https://www.researchgate.net/figure/Architecture-of-convolutional-neural-network-for-face-recognition_fig1_301703031

3.3.5. 3D ansiktsgenkänning

Metoder som Eigenface och Fisherface används för att analysera 2D bilder. Dessa metoder för 2D ansiktsgenkänning är dock opålitliga då det uppstår variation i belysning och olika vinklar på ansiktet. 3D ansiktsgenkänning skapar en "surface representation" av ansiktet. Den simplaste modellen för "surface representation" är ett "punktmoln", punkter som befinner sig i ett tredimensionellt plan, där varje punkt har x, y och z koordinater. Punktmolnets punkter kan jämföras med ett indata ansikte, där det euklidiska avståndet (avståndet mellan två punkter) jämförs.

4. Verktyg

I detta kapitel behandlas olika verktyg som kan användas för forskning inom ansiktsigenkänning. I nästa kapitel behandlas API:s för ansiktsigenkänning som används för kommersiellt bruk. För varje implementation kommer samma data set som inlärningsbilder. Verktygen kommer att utföra igenkänning och resultaten jämförs med tanke på prestanda och precision. Tanken är också att använda olika metoder för hur de metoderna som beskrevs i kapitel 3 implementeras i ett ansiktsigenkänningsprogram.

4.1. MATLAB

4.2. OpenCV

4.3. OpenFace

4.4. TensorFlow

5. APIs

I detta kapitel behandlas API:s för ansiktsgenkänning. Eftersom de flesta av API:na är kommersiella och avsedda för företag, kommer analysen av API:na att basera sig på en gratisversion/demo av systemet.

1. Amazon Rekognition
2. FaceX
3. Microsoft Face API
4. Anometrics Face Recognition
5. Face++
6. Google Cloud Vision
7. IBM Watson Visual Recognition

6. Diskussion

[I detta kapitel diskuteras resultaten från undersökningen av ansiktsgenkänningsverktyg och APIs. Resultaten kommer att analyseras på basen av pris, prestation, användning.]

7. Slutsats

[I detta kapitel sammanfattas avhandlingens innehåll och slutsatsen diskuteras på basen av analysen.]

8. Källor

- [1] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010.
- [2] A. C. Müller and S. Guido, *Introduction to Machine Learning with Python*, O'Reilly Media, 2017.
- [3] S. Z. Li and A. K. Jain, "Introduction," in *Handbook of Face Recognition*, Springer, 2011, pp. 1-18.
- [4] H. Liu, *Facial Detection and Recognition on Mobile Devices*, Morgan Kaufmann, 2015.
- [5] J.-K. Kämäräinen, A. Hadid and M. Pietikäinen, "Local Representation of Facial Features," in *Handbook of Facial Recognition*, Springer, 2011, pp. 79-108.
- [6] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, p. 16, 1991.
- [7] J. P. Hespanha, D. J. Kriegman and P. N. Belhumeur, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis & Machine Learning*, vol. 19, no. 7, pp. 711-720, 1997.
- [8] J. Patterson and A. Gibson, *Deep Learning*, O'Reilly Media, 2017.
- [9] D.-H. Kenneth, *A Practical Introduction to Computer Vision with OpenCV*, John Wiley & Sons, Incorporated, 2014.
- [10] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- [11] G. Bonaccorso, *Machine Learning Algorithms*, Packt Publishing, 2017.
- [12] M. Gori, *Machine Learning: A Constraint-Based Approach*, Morgan Kaufmann, 2017.
- [13] N. Buduma and N. Locascio, *Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms*, O'Reilly Media, 2017.
- [14] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*, O'Reilly Media, 2008.

- [15] S. Khan, H. Rahmani, S. A. A. Shah and M. Bennamoun, A Guide to Convolutional Neural Networks for Computer Vision, Morgan & Claypool Publishers, 2018.
- [16] H. Li, L. Li and K.-A. Toh, Advanced Topics in Biometrics, World Scientific, 2011.
- [17] T. Massimo, B. Josef and J. Anil, Biometric Authentication, Copenhagen: Springer, 2002.
- [18] V. J. R., Biometric Technologies and Verification Systems, Butterworth-Heinemann, 2007.
- [19] O. M. Parkhi, A. Vedaldi och A. Zisserman, Deep Face Recognition, 2015, British Machine Vision Conference.
- [20] A. Quaglia och C. M. Epifano, Face Recognition: Methods, Applications and Technology, Nova Science Publishers, 2012.