Karl Eriksson
2020-04-17 MT1a, Grupp 9 Antonio Ilestad

Miniprojekt 2: Klassificering av handskrivna siffror

1. Inledning

Allt fler branscher datoriseras. Ett exempel på detta är postkontoren. Posten är en bransch som, med hjälp av avläsningsprogram, effektiviserats avsevärt. En betydligt större mängd paket kan läsas av för att sedan sorteras och skickas gentemot att detta utförs av människor. En stor svårighet, i att läsa av siffror, är att alla har olika handstilar, vilket kan göra det svårt för datorer att läsa av siffror. Hur exakt kan ett program läsa av siffror?

2. Metod

För att kunna utföra projektet användes material från kursen vilket var en fil som heter datamängd och programvaran *Matlab*. Filen datamängd beskriver ett antal bilder av svarta handskrivna siffror mellan 0-9 på 16x16 pixlar från US postal service databas. Dessa blir alltså till 16x16 matriser där värdena i matriserna beskriver hur svart en viss pixel är. Filen består även av två delar, en träningsfas som har 7291 objekt och testmängd som har 2007 objekt.

För att kunna klassificera siffrorna från filen tilldelas varje siffra en egen klass vilket utförs i träningsfasen. Därefter kommer en testfas där man implementerar algoritmer som tilldelar okända siffror till rätt klass.

2.1 Tilldelning av klass

Klasserna för siffrorna skapas med formeln för singulärvärdesfaktorisering[1] [2].

$$A = U\Sigma V^t \qquad (1)$$

Då skapas en ny ON-bas matris för underrummet och rummet. Där A är en siffras matris, *U* är underrummets nya ortogonala matris, V^t är en ortogonal matris för rummet och Σ är en diagonalmatris med singulära värden. Efter skapandet av siffrornas underrum så testas de mot testdatan. Vilket görs genom att projicera på underrummen (2) och sedan räkna ut residualen (3).

$$x_i = BB^t * x \qquad (2)$$

x_i är projektionsvektorn där i:et är vilken siffra projektionen tillhör, B är ON-matrisen för underrummet och x är punkten som projiceras.

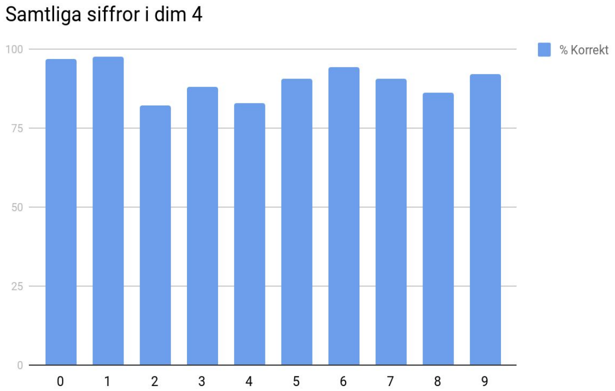
$$l = || x-x_i || \qquad (3)$$

Där x är den ursprungliga vektorn, x_i är projicerade vektorn och l är den normerade skillnaden mellan vektorerna. Samtliga 10 underrum projiceras på och siffran klassificeras som siffrans underrum som ger minst normerad residual. Efter att samtliga testobjekt har klassificerats som en siffra, jämförs resultatet med en svarslista som fanns i testdatan för att se hur effektiv algoritmen är.

---

[1] *Föreläsning 6&7 samt dess anteckningar TNA005*,Berkant, S ,Norrköping: LIU, 2020
[2] *Linjär algebra TNA002 för ED, KTS, MT,* Baravdish,G, Linköpings universitet, 2019

3.Resultat

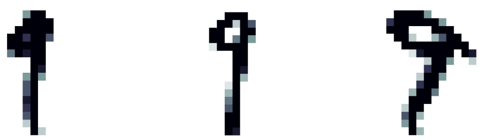


Figur 1: Jämförelse mellan implementerad algoritm och facit för samtliga siffror i procentform

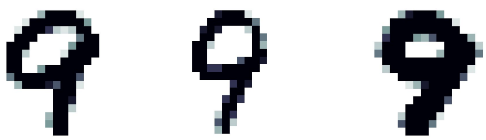
Dimension 4 är inte ett stort bevis på att algoritmen är ineffektiv. Vissa siffror kan ge svårigheter för algoritmen att klassificera i dim 4, dock togs dim 4 för att det är varken för lågt eller för högt, för att ge en viss översikt.

Dim:	4	6	8	10	12
Korrekt:	92,1%	92,7%	93,2%	92,7%	96,1%

Tabell 1: Jämförelse mellan implementerad algoritm och facit för siffran nio i olika dimensioner



Figur 2: Exempel på fel klassificerade nior



Figur 3: Exempel på rätt klassificerade nior

Siffran nio har tagits fram och analyserats extra. I figur 2 kan det ses att niorna saknar vissa egenskaper för nio, första nian har bland annat inget hål i sig och är för smal, vilket gör det svårt för algoritmen. När det kommer till rätt klassificerade nior, visar figur 3 att dessa nior är tydligt skrivna och har stora hål i sig och en tydlig svans, vilket behövs.

4. Diskussion

Det finns olika anledningar till att en siffra blir fel klassificerad vilket kan vara exempelvis att handskrivna siffror är slarvigt eller snabbt skrivna. När handskrivna siffror är närliggande varandra och ihoplänkade så blir det svårare att läsa av siffrorna, på grund av att blir svårare att veta var varje siffra börjar och slutar. Oftast handlar det även om att vissa siffror kan ha en viss likhet när det kommer till form, detta skapar problem för maskiner att då veta om en nolla är en nolla eller en sexa t.ex.

4.1 Slutsats

Projektion på underrum visar sig vara ett effektiv sätt att klassificera handskrivna siffror och det visar sig ha en bättre precision vid större dimensioner på underrummet som syns på tabel 1. I vissa lägen visar sig en lägre dimension vara mer precis än en högre, detta beror förmodligen på att algoritmen från början är väldigt precis och har svårt att få bättre resultat för varje dimension.

5. Referenser

1. Baravdish,G
   Linjär algebra TNA002 för ED,KTS,MT
   Linköpings universitet, 2019

2. Berkant, S
   Föreläsning 6&7 samt dess anteckningar TNA005
   Norrköping: LIU, 2020