



CHALMERS



Detektering, behandling och klustring av protoplanetära skivor i ALMA-arkivet

Utveckling av en oövervakad maskininlärningsmodell

Kandidatarbete i Rymd-, geo- och miljövetenskap

WILLIAM BERGQUIST, MARIE CARLANDER, THEA GRANSTRÖM,
CONRAD OLSSON, VILHELM NILSSON THORSSON & HANNES ÖHMAN

INSTITUTIONEN FÖR RYMD-, GEO- OCH MILJÖVETENSKAP

CHALMERS TEKNISKA HÖGSKOLA

Göteborg 2024

www.chalmers.se

KANDIDATARBETE 2024

Detektering, behandling och klustring av protoplanetära skivor i ALMA-arkivet

Utveckling av en oövervakad maskininlärningsmodell

WILLIAM BERGQUIST
MARIE CARLANDER
THEA GRANSTRÖM
CONRAD OLSSON
VILHELM NILSSON THORSSON
HANNES ÖHMAN



CHALMERS

Institutionen för Rymd-, geo- och miljövetenskap
CHALMERS TEKNISKA HÖGSKOLA
Göteborg 2024

Detektering, behandling och klustring av protoplanetära skivor i ALMA-arkivet
Utveckling av en oövervakad maskininlärningsmodell
WILLIAM BERGQUIST, MARIE CARLANDER, THEA GRANSTRÖM, CON-
RAD OLSSON, VILHELM NILSSON THORSSON & HANNES ÖHMAN

© William Bergquist, Marie Carlander, Thea Granström, Conrad Olsson, Vilhelm Nilsson Thorsson och Hannes Öhman, 2024.

Handledare: Per Bjerkeli och Maria Carmen Toribio Perez, Institutionen för Rymd-,
geo- och miljövetenskap
Examinator: Magnus Thomasson, Institutionen för Rymd-, geo- och miljövetenskap

Kandidatarbete 2024
Institutionen för Rymd-, geo- och miljövetenskap
Chalmers Tekniska Högskola
SE-412 96 Göteborg
Telefon +46 31 772 1000

Omslagsbild: En protoplanetär skiva runt HD 163296 från ALMA-arkivet [1].

Skriven i L^AT_EX
Göteborg 2024

Sammandrag

ALMA-observatoriet är ett av världens största observatorier med en enorm datamängd i sitt arkiv. För närvarande finns över 70 000 observationer i arkivet, på bland annat protoplanetära skivor. Forskningen kring skivorna fortskrider, men den stora datamängden försvårar manuell granskning av observationerna för att kunna dra nya insikter. Det är därför intressant att undersöka automatisk gruppering av datan. Maskininlärningsalgoritmer har visats sig värdefulla i liknande sammanhang. Detta projekt undersöker om algoritmerna även kan appliceras på ALMA-data.

Projektet implementerar en maskininlärningsmodell för automatisk klustring av observationer på protoplanetära skivor i ALMA-arkivet. Projektet avser att undersöka och utvärdera de mest effektiva metoderna inom bildbehandling och oövervakad maskininlärning. Den färdigställda modellen består av en bildbehandlings- och en maskininlärningsenhet.

Bildbehandlingsenheten syftar till att identifiera objekt och förbehandla bilder, exempelvis genom att minska brus och effekter av objektens position i bilden inför klustring med maskininlärningsenheten. Denna har visat sig vara kritisk för att förbereda bilder inför maskininlärningen som är uppdelad i tre delar: extrahering av attribut, dimensionsreducering och klustring. Olika algoritmer har testats och DCT, t-SNE och k -means med $k = 70$ kluster har visat sig vara lämpligast på respektive del för den valda datamängden.

Projektet har tagit fram en fungerande maskininlärningsmodell som underlättar arbetet att hitta likartade protoplanetära skivor. Det här projektet förmodas bidra till att främja automatiserade metoder för analys av omfattande astronomiska data. Vidare kan det öka förståelsen för protoplanetära skivor och deras roll i bildandet av planetsystem.

Nyckelord: protoplanetära skivor, ALMA, oövervakad maskininlärning, klustring, bildbehandling, dimensionsreducering.

Abstract

The ALMA observatory is one of the world's largest observatories with an enormous amount of data in its archive. There are currently over 70,000 observations in the archive, including observations of protoplanetary disks. The research into the disks continues, but the large volume of data poses challenges for manual data analysis. Machine learning algorithms have proven valuable in similar contexts. The project investigates whether these algorithms also can be applied to ALMA data.

This project aims to implement a machine learning model for automated clustering of observations in the ALMA archive. To achieve this, the project investigates and evaluates the most effective methods in image processing and unsupervised machine learning. The developed model consists of an image processing unit and a machine learning unit.

The image processing unit aims to identify stars and preprocess images in order to reduce noise and the effects of object positioning for the following clustering with the machine learning unit. Image processing proved to be critical for correctly clustering the dataset in an unsupervised manner. The machine learning unit has analysed many different algorithms and is divided into three parts: attribute extraction, dimensionality reduction, and clustering. The model that has been proven most prominent in this project is the combination of feature extraction with DCT, dimensionality reduction with t-SNE, and clustering with k -means with $k = 70$ clusters.

The project has developed a functioning machine learning model that facilitates the task of finding similar stars and protoplanetary disks. This project is expected to contribute to advancing automated methods for analyzing large-scale astronomical data. Furthermore, the project may enhance the understanding of protoplanetary disks and their role in system formation.

Keywords: protoplanetary disks, ALMA, unsupervised machine learning, clustering, image processing, dimensionality reduction.

Förord

Vi vill framföra ett varmt tack till Per Bjerkeli och Maria Carmen Toribio Perez för deras ovärderliga stöd som handledare i detta projekt.

Per, med sin stora entusiasm och sin lättsamma humor, har varit en fantastisk ledare och ett stöd för oss genom hela processen. Hans brinnande intresse för AI och expertis inom astronomi har inspirerat oss och drivit projektet framåt.

Carmen har bidragit med enorma kunskaper om ALMA-observatoriet och säkerställt att vi hade tillgång till rätt resurser och bakgrundsfakta för att maximera möjligheten att lyckas. Hennes professionalism och noggrannhet har varit av avgörande betydelse för projektets framgång.

Vi är evigt tacksamma för deras dedikation och stöd.

William Bergquist, Marie Carlander, Thea Granström, Conrad Olsson,
Vilhelm Nilsson Thorsson & Hannes Öhman
Göteborg, maj 2024

Akronymer

Nedanför finns en lista med akronymer som använts i rapporten listade i alfabetisk ordning.

ACA	Atacama Compact Array
ALMA	Atacama Large Millimeter/submillimeter Array
BIC	Bayesian Information Criterion
CNN	Faltningsnätverk
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DCT	Diskret Cosinustransform
FITS	Flexible Image Transport System
GMM	Gaussian Mixture Model
HOG	Histogram of Oriented Gradient
t-SNE	t-distributed Stochastic Neighbor Embedding
WCSS	Within Cluster Sum of Squares
PCA	Principal Component Analysis

Innehåll

Akronymer	x
1 Inledning	1
1.1 Syfte	2
1.2 Avgränsningar	2
2 Teori	3
2.1 Protoplanetära skivor	3
2.2 ALMA-observatoriet	3
2.2.1 ALminer	4
2.2.2 Flexible Image Transport System	4
2.2.3 Filnamnstillägg	5
2.3 Oövervakad Maskininlärning	5
2.3.1 Extrahering av bildattribut	6
2.3.1.1 Pixelvärden	6
2.3.1.2 Histogram of Oriented Gradients	7
2.3.1.3 Diskret cosinustransform	7
2.3.1.4 Faltningsnätverk	8
2.3.2 Dimensionsreducering	9
2.3.2.1 t-Distributed Stochastic Neighbor Embedding	9
2.3.2.2 Principalkomponentanalys	9
2.3.3 Klusteranalys	10
2.3.3.1 K-means	10
2.3.3.2 Gaussian Mixture Model	11
2.3.3.3 DBSCAN	12
2.3.4 Silhuettvärden	13
2.3.5 Scikit-learn	14
3 Metod	17
3.1 Datahämtning	17
3.1.1 Filtrering	17
3.2 Bildbehandling	18
3.2.1 Beskärning av bild	18
3.2.2 Selektiv bildeliminering	18
3.2.3 Förstoring	19
3.2.4 Maskering med oskärpa	20

3.2.5	Skaländring och rotation	21
3.2.6	Sparande av bilddata	24
3.2.7	Omvandling till JPG	24
3.3	Maskininlärning	25
3.3.1	Extrahering av bildattribut	25
3.3.2	Dimensionsreducering	25
3.3.3	Klustring	25
3.4	Utvärdering	26
4	Resultat	27
4.1	Datahämtning och brusfiltrering	27
4.2	Bildprocessering	27
4.2.1	Rotation	29
4.3	Maskininlärningsmodeller	30
4.3.1	Jämförelse av attributextraheringsmodeller	31
4.3.2	Dimensionsreducering med t-SNE	33
4.3.3	Klusterallokering med k -means	34
4.3.4	Klusterallokering med DBSCAN	35
4.3.5	Utvärdering av tester	37
5	Diskussion	39
5.1	Datahämtning	39
5.2	Bildprocessering	39
5.2.1	Beskärning	39
5.2.2	Brusreducering	40
5.2.3	Selektiv bildeliminering	40
5.2.4	Förstoring	41
5.2.5	Skaländring och rotation	42
5.3	Maskininlärning	43
5.3.1	Attributextrahering	43
5.3.2	Dimensionsreducering	44
5.3.3	Klustringsmodeller	45
5.4	Jämförelse med liknande projekt	46
6	Slutsats	49
	Bibliography	51
A	Övriga resultatbilder	I
B	Klusterinformation	IX
C	Kod	XVII
D	Bildmanipuleringsinformation	XIX

1

Inledning

Atacama Large Millimeter/submillimeter Array (ALMA), som är beläget i de norra delarna av Anderna i Chile, räknas som ett av världens största observatorier [2]. Med sin kontinuerliga datainsamling växer ALMA-arkivet dagligen och innehåller för närvarande närmare 70 000 unika observationer [3]. Denna enorma mängd data gör det svårt att manuellt granska varje observation, vilket ökar risken för att viktiga upptäckter förbises. Därför är automatisering av bildgruppering i arkivet av stort intresse.

Enligt Ronald D Ekers [4] var tio av sjutton stora astronomiska upptäckter under 1900-talet resultatet av oavsiktliga iakttagelser snarare än avsiktliga studier. I och med detta poängteras vikten av att implementera objektiva klustringsystem i ALMA-arkivet, vilket kan leda till nya vetenskapliga upptäckter. Ett exempel på ett område där detta har relevans är planetbildning.

En typ av objekt som ALMA kan observera är protoplanetära skivor, som består av gas och stoft kring nybildade stjärnor. De spelar en central roll i både bildandet och utvecklingen av stjärnor och planetsystem [5]. Rollen skivorna spelar när systemen bildas är inte helt klar. Enskilda studier bidrar bara med en del av helheten och för en djupare förståelse krävs analys av stora mängder data. Eftersom manuell inspektion kräver fler resurser än vad en forskargrupp besitter, behövs förbättrade och automatiserade metoder.

Tidigare studier, till exempel arbetet av Fagrell, Kollberg, Rasmussen m. fl. [6], har belyst hur maskininlärningstekniker kan utnyttjas för att klassificera och analysera astronomiska bilder. Arbetet fokuserade på övervakad maskininlärning och avancerade metoder för bildanalys för att gruppera bilder från ALMA-arkivet. Även om metoderna är annorlunda, inspirerar de direkt de tekniker som används i detta projekt.

I det här projektet valdes istället att utforska hur oövervakad maskininlärning kan bidra till att gruppera observationer. Denna metod skiljer sig från den tidigare nämnda övervakade maskininlärningen, som använder förhandsmärkt data för att träna en modell. Som kontrast till detta tillåter oövervakad maskininlärning att objektivet utforska data utan förutbestämda etiketter, vilket är avgörande för att upptäcka nya och tidigare oidentifierade fenomen, exempelvis inom astronomi.

1.1 Syfte

Projektets syfte är att utveckla och implementera en maskininlärningsmodell för att effektivt gruppera observationer från det offentligt tillgängliga ALMA-arkivet, med ett särskilt fokus på protoplanetära skivor. I arkivet finns för närvarande närmare 6 600 observationer av protoplanetära skivor, vilket ger möjlighet att objektivt gruppera och analysera data [3].

Genom att automatisera klustring på bilder av dessa skivor, syftar modellen till att förenkla navigeringen i arkivet och förbättra tillgängligheten samt användbarheten av datan. Detta kan underlätta identifieringen av relevanta observationer och möjliggöra statistiska slutsatser om de olika grupperna och egenskaperna hos skivorna. Mer specifikt syftar projektet till att identifiera och utvärdera de mest effektiva metoderna för bildbehandling och oövervakad maskininläring för att klustra astronomiska bilder.

1.2 Avgränsningar

I detta projekt beslutades att endast behandla tvådimensionella bilder, trots att ALMA-observatoriet har kapacitet att ta bilder vid hög frekvensupplösning vilket ger tredimensionella datakuber. Denna avgränsning minskar komplexiteten i maskininlärningsmodellen, vilket gör utvecklingen av en effektiv modell möjlig. På grund av ALMAs omfattande databas valdes även att specifikt kategorisera observationer av protoplanetära skivor. Denna ytterligare avgränsning minskar den datamängd som behöver hanteras, vilket ökar sannolikheten för att modellen ska fungera som väntat och förenklar bildbehandlingen.

2

Teori

I detta kapitel presenteras den teoretiska grund som är nödvändig för rapporten. Kapitlet täcker de astronomiska principer som projektet bygger på och beskriver de algoritmer som har använts.

2.1 Protoplanetära skivor

Protoplanetära skivor spelar en central roll i bildandet av planetsystem [5]. Dessa skivor bildas som en naturlig följd av bevarandet av rörelsemängdsmoment under kollapsen av protostellära moln och omger den nybildade stjärnan. Genom skivorna kan planetbildning ske genom flera processer. En möjlig process är ackretion där stoftpartiklar gradvis klumpar ihop sig för att bilda växande fasta objekt, från små stoftkorn till planetesimaler och vidare till fullskaliga planeter [7]. Alternativt kan delar av skivan genomgå en gravitationell kollaps, vilket leder till bildandet av större kroppar. Skivans densitet och temperaturprofiler samt dess storlek och massa är avgörande för de processerna.

Enligt Lebreuilly, Hennebelle, Colman m. fl. [5] har den teoretiska forskningen kring protoplanetära skivor utvecklats avsevärt under de senaste två decennierna. Trots framstegen kvarstår dock signifikanta osäkerheter gällande de initiala egenskaperna hos skivorna. Lebreuilly, Hennebelle, Colman m. fl. [5] menar att observationer av protoplanetära skivor i tidiga stadier är begränsade, eftersom skivorna är mindre och fasen varar under en begränsad tid. Dessutom är de tidiga skivorna ofta djupt inbäddade i tjocka stjärnmoln, vilket komplicerar direkt observation.

2.2 ALMA-observatoriet

ALMA är världens största astronomiska observatorieanläggning, belägen på Chajnantorplatån i Anderna i norra Chile [8]. ALMA-projektet inleddes som ett internationellt samarbete mellan flera institutioner: National Radio Astronomy Observatory, förvaltad av Associated Universities, Inc., som representerar Nordamerika; National Astronomical Observatory of Japan för Östasiens räkning; och Europeiska södra observatoriet [9].

Observatoriet består av 66 högprecisionsantennor som arbetar över våglängder från 0,32 mm till 3,6 mm [9]. Genom användning av interferometriteknik integreras dessa antenner för att fungera som ett enda stort radioteleskop där huvuddelen av

anläggningen, Main Array, omfattar 50 stycken 12-metersantennor. Dessutom finns Atacama Compact Array (ACA), som inkluderar fyra 12-metersantennor och tolv 7-metersantennor. Main Array och ACA bildar tillsammans ALMA. Antennornas konfigurationer kan anpassas där avstånden mellan dem varierar från 150 meter upp till 16 kilometer, vilket är avgörande för den slutliga bildens spatiala upplösning.

Data som samlas in av ALMA genom interferometri motsvarar fouriertransformen av den observerade regionen [10]. För att rekonstruera en bild från data används ofta CLEAN-algoritmen, utvecklad av Högbom [11], vilket är en väletablerad metod för att hantera de spår av interferensmönster som kan uppstå i bilderna. Denna process har, trots sin komplexitet, effektiviserats och automatiserats, vilket möjliggör att ALMA kan tillhandahålla bearbetade astronomiska bilder i arkivet [12]. Detta möjliggör för användare att dra nytta av högkvalitativ data utan att själva behöva hantera de mest krävande aspekterna av datareducering.

2.2.1 ALminer

ALminer är ett datautvinningsverktyg utvecklat genom ett samarbete mellan Allegro, ALMA Regional Centre i Nederländerna och University of Vienna [13]. Verktöget är utformat för att underlätta nedladdning, analys och visualisering av data från ALMA-arkivet. Med en Python-baserad kodstruktur erbjuder ALminer användarna möjligheten att effektivt filtrera och söka data baserat på nyckelord från arkivets metadata. Denna funktionalitet möjliggör snabbt identifiering av relevanta observationer och positioner.

ALminer tillhandahåller sin väldokumenterade kod öppet via GitHub, vilket underlättar för användare att anpassa verktyget efter sina egna behov [13]. Genom att erbjuda tillgång till både behandlad data och rådata möjliggör ALminer ett flexibelt sätt att utforska och analysera ALMA-data. Verktöget kompletteras med en detaljerad användarmanual, tillgänglig på deras webbplats, vilket ger ytterligare stöd och vägledning för att maximera dess potential för studier.

2.2.2 Flexible Image Transport System

Flexible Image Transport System (FITS) är ett standardiserat format som används globalt för att överföra astronomiska data. Detta format, som har erkänts av NASA och den Internationella Astronomiska Unionen, utvecklades för att möjliggöra en enhetlig metod för dataöverföring [14]. Till skillnad från vanliga bildformat som JPG eller GIF, erbjuder FITS en exceptionell mångsidighet och används inte enbart för bilder utan också för transport, analys och arkivering av olika vetenskapliga datamängder.

Vidare stöder FITS flera dimensioner, inklusive 1D-spektra, 2D-bilder och 3D-datakuber, samt tabeller. Formatet innehåller även nyckelord som ger beskrivande information om data och stöder moderna komprimeringsmetoder för effektiv lagring och över-

föring. Principen “*once FITS, always FITS*” står som grund för FITS-standardens stabilitet och innebär att befintliga FITS-filer inte får modifieras på ett sätt som gör dem inkompatibla med den etablerade standarden [15]. Alla förändringar och tillägg till FITS-standard granskas noggrant av internationella astronomiska organisationer innan de implementeras, vilket säkerhetsställer enhetlighet och korrekt tolkning av data inom astronomin.

2.2.3 Filnamnstillägg

I ALMA-arkivet används ett förutbestämt schema för att namnge FITS-filer, vilket underlättar identifiering och hantering av data. Filnamnstillägget “*_sci*” betyder att bilden är tagen för ett vetenskapligt syfte, snarare än för kalibrering eller andra kontroller [16]. Vidare representerar tillägget “*.cont*” kontinuumbilder, vilka består av data där spektrallinjer har filtrerats bort för att framhäva kontinuumemissionen, oftast sammansatt från data över ett brett spektrum av våglängder. Denna representation är fördelaktig för att studera egenskaper av objekt i rymden utan störning från specifika linjedata.

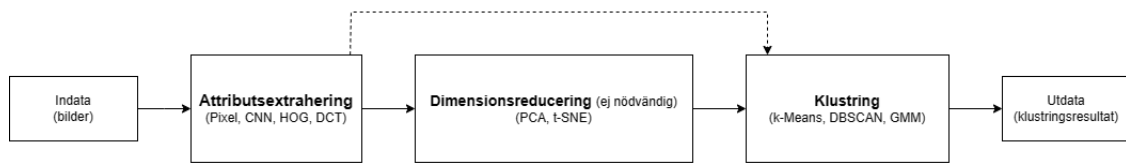
Slutligen refererar tillägget “*.pbcor*” till en korrigerings som gjorts för att kompensera för variationer i känslighet över en ALMA-antenns synfält, vilket kallas huvudloben [17]. Eftersom antennens känslighet är högst i mitten och minskar ut mot kanterna, säkerställer denna korrigerings att ett objekt framstår med samma intensitet oavsett dess placering i bilden. Detta innebär dock att brusnivån kommer att variera med avståndet från bildens centralpunkt och orsaka brus med starkare kontrast närmare kanterna. Korrigerings genomförs genom att dividera de observerade värdena med observationens motsvarande “*.pb*” fil som representerar den radiella känsligheten hos antennen.

2.3 Övervakad Maskininlärning

Övervakad maskininlärning representerar en kategori inom maskininlärning som fokuserar på att analysera datamängder utan fördefinierade etiketter eller strukturer. Denna teknik använder algoritmer för att identifiera distinkta mönster, grupperingar och likheter, helt utan mänsklig interaktion. Enligt IBM [18] ligger styrkan i oövervakad maskininlärning i dess förmåga att hantera ostrukturerade och komplexa datamängder, vilket gör det möjligt att identifiera samband och likheter som annars skulle förbli oupptäckta.

Ett av de mest framstående användningsområdena för oövervakad maskininlärning är klusteranalys. Målet med klusteranalys är att organisera objekt i grupper, eller kluster, så att objekten inom varje grupp är mer lika varandra än objekten i andra grupper. Denna metod är särskilt användbar inom områden där data inte naturligt kopplas till etiketter, eller där de underliggande kategorierna inte är kända i förväg.

Processen för att genomföra en klusteranalys med oövervakad maskininlärning kan



Figur 2.1: Diagram som beskriver uppbyggnaden av en klustringsmodell enligt de tre huvudsakliga stegen detta projekt arbetar efter.

struktureras enligt illustrationen i figur 2.1. Här används en modell där data initialt genomgår attributextrahering, följt av dimensionsreducering för att förenkla data innan den slutliga klusteranalysen utförs. Kortfattat består varje steg av:

- **Attributextrahering:** Det första steget där rådata omvandlas till en form som kan bearbetas av maskininlärningsmodellen. Exempel på tekniker som kan extrahera relevant information är pixelvärden, faltningsnätverk (CNN), histogram of oriented gradients (HOG) och diskret cosinustransform (DCT).
- **Dimensionsreducering:** Detta steg är avgörande för att hantera de stora mängder data som genereras i föregående steg. Tekniker som principal component analysis (PCA) och t-distributed stochastic neighbor embedding (t-SNE) kan testas för att minska dimensionerna av data, vilket behåller de mest signifikanta egenskaperna samtidigt som det reducerar komplexiteten.
- **Klusteranalys:** Slutligen används klusteranalysmetoder såsom exempelvis k -means, DBSCAN och Gaussian mixture models (GMM) för att gruppera data baserat på de extraherade och reducerade egenskaperna. Genom att identifiera och gruppera liknande datapunkter kan mönster upptäckas och tolkas.

2.3.1 Extrahering av bildattribut

Extrahering av bildattribut utgör det initiala steget i maskininlärningsprocessen, där varje bild omvandlas till en attributvektor [19]. Denna vektor innehåller kritiska egenskaper från bilden, anpassade för effektiv bearbetning av maskininlärningsalgoritmer. Beroende på bildens karaktär och struktur kan olika metoder för attributextrahering vara mer eller mindre lämpliga. För detta projekt har fyra olika tekniker testats, var och en anpassad för att möta specifika krav och variationer i data:

- **Pixelvärden**
- **Histogram of Oriented Gradients (HOG)**
- **Diskret Cosinustransform (DCT)**
- **Faltningsnätverk (CNN)**

2.3.1.1 Pixelvärden

Pixelvärden utgör det kanske mest intuitiva sättet att extrahera attribut från en bild. I detta projekt behandlas svartvita bilder, där ljusstyrkan i varje pixel används som ett attribut [20]. För en bild med dimensionerna $M \times N$ genereras en attributvektor bestående av $M \times N$ element. Denna process innebär att bilden, som representeras av en matris, omformas till en vektor genom att raderna läggs efter varandra i en sekventiell ordning.

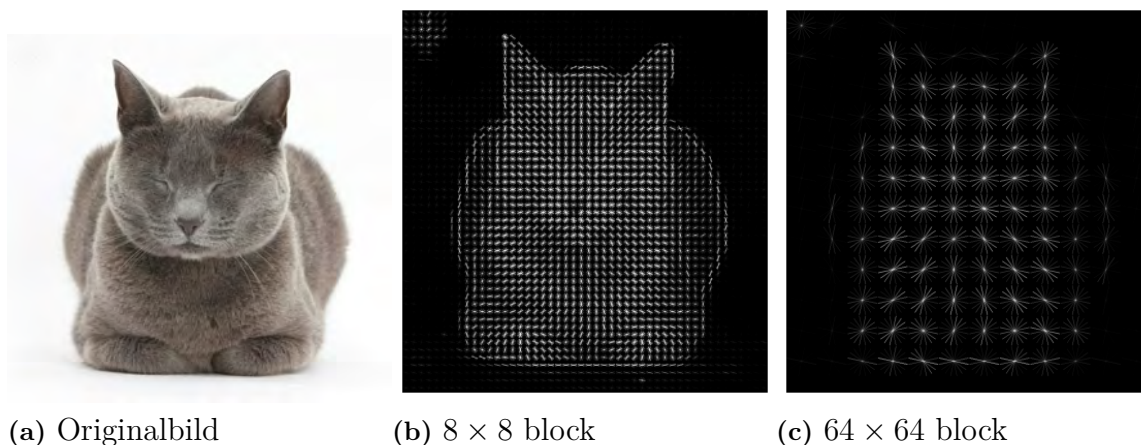
2.3.1.2 Histogram of Oriented Gradients

HOG är en annan algoritm som används för att extrahera attribut från bilder [21]. Algoritmen delar in bilden i block där varje block kan delas in i flera celler. För varje pixel p_i i cellen beräknas gradienterna G_x, G_y i x - och y -riktningarna. Gradienternas magnitud r och riktning θ beräknas för varje pixel med formlerna

$$r = \sqrt{G_x^2 + G_y^2}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right).$$

I varje cell skapas ett histogram baserat på intervall av vinklar i intervallet $[0, 2\pi)$. Riktningsektorn för varje vinkel beräknas utifrån pixelns gradienter. Cellernas histogram, som visar fördelningen av riktningsektorer av olika storlekar, lagras sedan som attribut i en vektor. Metodens förmåga att identifiera kanter i bilder är dess största fördel, eftersom områden med skilda ljusstyrkor skapar höga värden på gradienten i de riktningar som är vinkelräta mot kanterna. Ett exempel på hur detta kan se ut för olika storlekar på block illustreras i figur 2.2.



Figur 2.2: Exempel på hur HOG-algoritmen transformerar bilder för två olika blockstorlekar [22].

2.3.1.3 Diskret cosinustransform

DCT är en metod som liknar pixelvärdemetoden som beskrivs i avsnitt 2.3.1.1. Denna metod utför dock en diskret cosinustransform på alla pixelvärden i en bild, varpå den resulterande matrisen används som grund för en attributvektor.

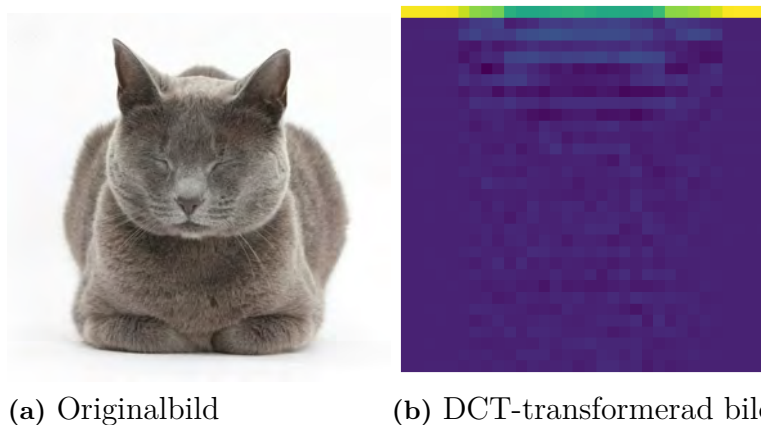
Antag att bilden är svartvit med dimensionen $M \times N$. Låt $p(x, y)$ beteckna ljusstyrkan i pixeln med position (x, y) . Den diskreta cosinustransformen av ljusstyrkan,

$P(u, v)$, bestäms enligt följande formel

$$P(u, v) = \frac{1}{\sqrt{MN}} \alpha(u) \alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} p(x, y) \cos\left(\frac{(2x+1)u\pi}{2M}\right) \cos\left(\frac{(2y+1)v\pi}{2N}\right)$$

$$\text{där } u = 1, 2, \dots, N, \quad v = 1, 2, \dots, M, \quad \alpha(k) = \begin{cases} 1/\sqrt{2} & \text{om } k = 0 \\ 1 & \text{annars} \end{cases}.$$

Denna metod tillåter en effektiv komprimering av bildens frekvensdata genom att omvandla den råa bildmatrisen till en komprimerad form. Efter att ha tillämpat DCT transformeras bildmatrisen till en vektor av längd $M \times N$. De viktigaste beståndsdelarna av dessa data finns då i början av attributvektorn, vilket möjliggör att vektorn kan kapas utan lika stora förluster av information. Detta är speciellt relevant då bilderna har hög dimension för att minska mängden attribut som används i senare steg av maskininlärningsmodellen. Ett exempel på en DCT-transformerad bild visas i figur 2.3.



Figur 2.3: Exempel på hur DCT-algoritmen transformerar bilder [22].

2.3.1.4 Faltningsnätverk

Faltningsnätverk, även känt som CNN, är en speciell typ av artificiella neuronnät som är särskilt lämpade för att hantera bild- och ljudsignaler [23]. CNN består av flera lager av noder, varav de dolda lagren främst utgörs av faltningslager som är designade för att effektivt extrahera attribut från bilder.

Ett CNN efterliknar den mänskliga hjärnans förmåga att känna igen mönster genom att först använda ett indata-lager som tar emot den ursprungliga datamängden [24]. Det följs av flera dolda lager som utför olika uppgifter. I de dolda lagren identifieras och extraheras mönster och funktioner från indata genom en process kallad faltning. Varje lager transformeras vidare för att fokusera på specifika aspekter av denna data, vilket möjliggör en gradvis och effektiv inläring av komplexa strukturer.

Det yttersta lagret i CNN ger slutresultatet av nätverkets bearbetning [25, s.21-22]. I tillämpningar där CNN används för attributextrahering avlägsnas detta sista la-

ger vanligtvis och det näst sista lagret betraktas då som en attributvektor. Denna vektor representerar de viktigaste funktionerna som identifierats från bilderna och används för vidare analys.

Teorin om faltningsnätverk är omfattande och detta avsnitt är endast en översiktlig introduktion till hur de används för att extrahera värdefulla attribut från bilder inom ramen för maskininlärningsbaserade klustringsmodeller. Ytterligare information om ämnet finns tillgänglig i O'Shea och Nash [26], som ger en bredare översikt.

2.3.2 Dimensionsreducering

Dimensionsreducering är ett steg i maskininlärningsmodellen vars syfte är att reducera dimensionen av attributvektorn med minimal förlust av information [27]. Längden på attributvektorn kan variera beroende på indata och den valda attributextraheringsmetoden, vilket kan resultera i stora vektorer som ökar beräkningstiden. För att effektivisera beräkningarna krävs ofta en reduktion av vektorlängden. Att selektivt betrakta endast en del av vektorn är en strategi, men detta medför en betydande förlust av attribut och kan resultera i försämrade klustring. Det som behövs är en metod för att behålla den mest betydelsefulla informationen från attributvektorn och förkasta överskottet. Detta kan uppnås på flera sätt och i detta arbete har två algoritmer implementerats vilka presenteras i kommande avsnitt.

2.3.2.1 t-Distributed Stochastic Neighbor Embedding

Metoden t-distributed stochastic neighbor embedding (t-SNE) är ett verktyg inom maskininlärning och dataanalys som kan användas som förbehandlingssteg till andra algoritmer. Den slutgiltiga metoden utvecklades av Laurens van der Maaten som lade till "t-" i namnet på den ursprungliga "SNE" metoden som Geoffrey Hinton och Sam Roweis utvecklade [28].

Syftet med algoritmen är att bevara den lokala strukturen hos datapunkter i den högdimensionella rummet när de projiceras ner till en lägre dimensionell rum, vanligtvis 2D eller 3D, för att underlätta visualisering. Algoritmen skapar en sannolikhetsdistribution som representerar likheterna mellan datapunkter i den högdimensionella rummet. Därefter försöker den minimera skillnaderna mellan denna distribution och en motsvarande distribution i den lägre dimensionen. På så sätt kan den reducera dimensionen för att visualisera högdimensionell data på ett effektivt och begripligt sätt.

2.3.2.2 Principalkomponentanalys

Principalkomponentanalys (PCA) är en statistisk metod som syftar till att reducera dimensionerna inom en datamängd och samtidigt minimera förlusten av viktig information [18]. Det primära målet med PCA är att identifiera mönster i data genom att transformera de ursprungliga variablerna till en ny uppsättning av oberoende variabler, kända som huvudkomponenter. Komponenterna arrangeras så att de fångar upp den största möjliga variansen i datamängden med de första komponenterna och

därefter avtar betydelsen av de efterföljande komponenterna i fallande ordning.

Genom att fokusera på de mest signifikanta variationerna tillåter PCA en effektiv datakomprimering. Detta görs genom att reducerade dimensioner fortfarande bevarar essensen av de ursprungliga datamängderna. På detta vis kan antalet variabler som behövs för att adekvat beskriva datamängden reduceras. Denna reduktion är särskilt värdefull i förberedande steg för mönsterigenkänning och klusteranalys, där överskådlighet och bearbetbarhet i datamängden kan förbättra analysens noggrannhet och effektivitet.

2.3.3 Klusteranalys

Klusteranalys är en metod inom oövervakad maskininlärning vars mål är att gruppera ostrukturerad data baserat på deras likheter eller skillnader [18]. Genom användning av klusteralgoritmer bearbetas råa och omärkta dataobjekt för att organiseras enligt strukturer eller mönster.

2.3.3.1 *K*-means

Exklusiv klusterallokering är en typ av gruppering som fastställer att varje dataobjekt endast kan tillhöra ett kluster. Enligt IBM [18] är en allmänt använd algoritm inom detta område *k*-means. Den grundläggande principen för *k*-means är att minimera within-cluster sum of squares (WCSS), vilket är summan av variansen mellan objekt i varje kluster. Detta görs genom att hitta en uppdelning av datamängden $C = \{c_1, \dots, c_k\}$ så att summan av kvadrerade avstånd från varje punkt till dess närmaste centrum är minimal [29]. Minsta WCSS kan enligt Kriegel, Schubert och Zimek [29] beräknas genom att använda ekvationerna

$$WCSS_1 = \sum_{\substack{c_i \in C \\ \text{Kluster}}} \sum_{j=1 \dots d} \sum_{x, y \in c_i} (x_{ij} - y_{ij})^2 \quad (2.1)$$

$$WCSS_2 = \sum_{\substack{c_i \in C \\ \text{Kluster}}} \sum_{j=1 \dots d} 2|c_i| \sum_{\substack{x \in c_i \\ \text{Objekt}}} (x_{ij} - \mu_{ij})^2, \quad (2.2)$$

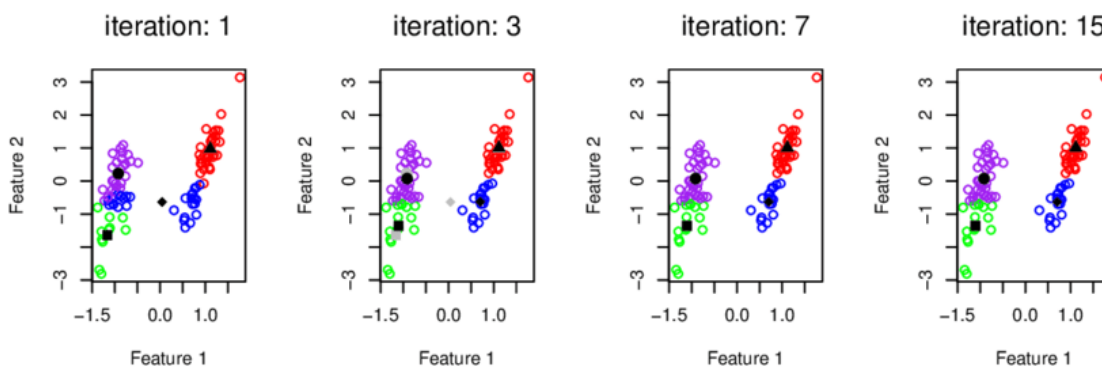
där författaren menar att μ_{ij} är medelkoordinaten för *i*:te klustret i *j*-dimensionen. Ekvation 2.1 representerar summan av parvisa avstånd inom varje kluster, vilket innebär att den beräknar den totala variationen eller spridningen av datapunkter inom varje kluster. Emellertid omformulerar ekvation 2.2 detta till att beräkna summan av klustervarianser, där varje klustervarians är viktad efter klusterstorlek. Detta implicerar att större kluster kommer att utöva en proportionellt större effekt på den totala variationen.

Enligt Kriegel indikerar båda ekvationerna att målet är att minimera kvadrerade euklidiska avstånd, antingen genom att minimera avstånden mellan klustermedlemmar och deras centrum eller genom att minimera det totala kvadrerade avståndet

för alla datapunkter.

Även om det är en allmän uppfattning att k -means optimerar den totala summan av euklidiska avstånd, menar Krigel att detta inte fallet. Den totala summan av euklidiska avstånd förblir konstant under k -means iterationer. Istället bryts totalvariationen av datamängden ned i två delar: WCSS och mellan-kluster-summa av kvadrater. Minimering av WCSS innebär att öka avståndet mellan objekt som finns i olika kluster, vilket i slutändan leder till en bättre klusteruppdelning.

Under varje iteration av k -means flyttas klustercentrarna för att minska WCSS. Detta innebär att centroiderna justeras för att minska avstånden mellan datapunkterna och deras tillhörande centroider vilket demonstreras i figur 2.4



Figur 2.4: Exempel på k -means iterationer från Rodriguez, Comin, Casanova m. fl. [30]. Klustercentrarna är markerade med svart, medan datapunkterna är färgkodade enligt deras tilldelade kluster. Grå markörer visar var klustercentrarna låg i den föregående iterationen. Även om datamängden endast innehåller 2 kluster användes $k = 4$ i algoritmen.

2.3.3.2 Gaussian Mixture Model

IBM [18] beskriver även en annan typ av klusteranalys vilket är en probabilistisk modell. Denna grupperar datapunkter baserat på deras sannolikhet att tillhöra en specifik fördelning, vilket hanterar densitetsuppskattning och mjuka klusterproblem. Gaussian Mixture Model (GMM) är en välkänd metod inom detta område. GMM illustreras av Reynolds [31] som en viktad summa av M komponenter av Gaussiska densiteter enligt

$$p(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \sum_{i=1}^M \gamma_i g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \quad (2.3)$$

där \mathbf{x} representerar en D -dimensionell datavektor med kontinuerliga värden, exempelvis mätningar eller egenskaper. γ_i är blandningsvikten och de komponent-Gaussiska densiteterna är $g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$. Vidare är varje komponentdensitet en Gaussisk

funktion med D variabler enligt

$$g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)' \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right\}, \quad (2.4)$$

där $\boldsymbol{\mu}_i$ är medelvektorn och $\boldsymbol{\Sigma}_i$ är kovariansmatrisen.

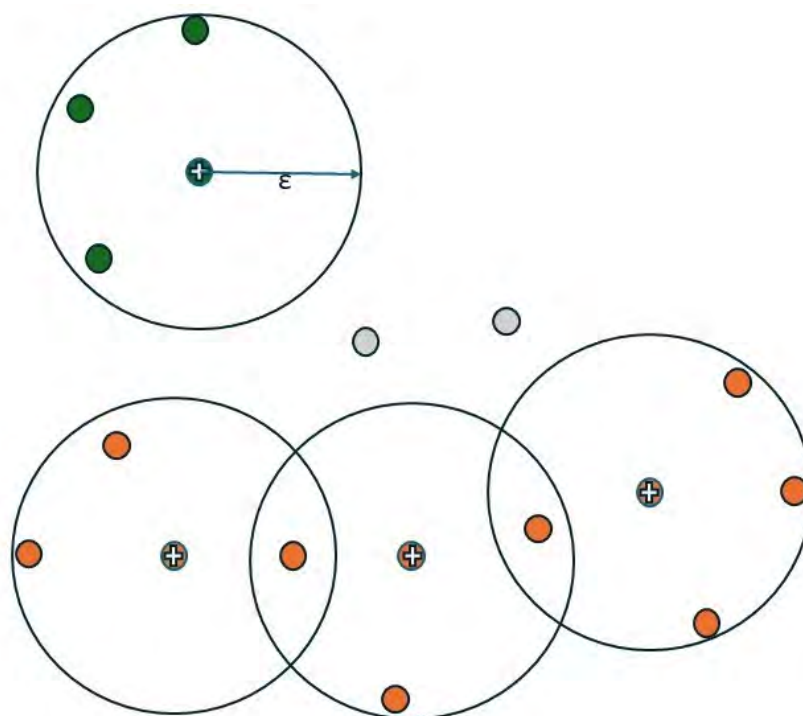
2.3.3.3 DBSCAN

En annan algoritm som är vanlig inom klusteranalys är Density-Based Spatial Clustering of Applications with Noise (DBSCAN), vilket är en algoritm för att upptäcka kluster i stora spatiala databaser med brus [32]. DBSCAN skiljer sig från andra klustringsalgoritmer genom dess förmåga att hantera brus och identifiera kluster baserade på täthet snarare än avstånd.

DBSCAN kategoriserar punkter i en av tre typer: kärnpunkter, gränspunkter eller bruspunkter. En kärnpunkt har ett visst antal punkter inom ett definierat avstånd, ε (epsilon) och dessa punkter utgör en del av ett kluster. En gränspunkt ligger inom avståndet ε från en kärnpunkt men uppfyller inte kravet på minst antal punkter för att vara en kärnpunkt. En bruspunkt är inte en del av ett kluster.

DBSCAN-algoritmen använder två centrala begrepp: ε och “*min_samples*”. ε definierar ett område runt varje punkt där andra punkter anses vara grannar om de ligger inom detta avstånd. “*min_samples*” är det minsta antalet punkter som måste finnas inom ε -avståndet för att en punkt ska betraktas som en kärnpunkt.

En kärnpunkt är en punkt som har åtminstone “*min_samples*” grannar, inklusive sig själv, inom dess ε -område. Algoritmen börjar med en slumpmässig punkt och identifierar sedan alla punkter som är direkt täthetsnära till denna kärnpunkt, vilket bildar ett kluster. Denna process upprepas rekursivt för att inkludera alla punkter som tillhör samma kluster. Punkter som varken är kärn- eller gränspunkter betraktas som bruspunkter. Dessa koncept illustreras i figur 2.5.



Figur 2.5: Exempel på två kluster, grön respektive orange, i DBSCAN där “*min_samples*” är satt till fyra. Punkterna med ett kryss är kärnpunkterna, övriga orangea samt gröna är gränspunkter och de gråa är bruspunkter.

DBSCAN:s förmåga att hantera olika tätheter och brus gör den särskilt användbar för spatiala datamängder där klustrens form och storlek kan variera avsevärt. Dess effektivitet även på stora datamängder och förmågan att identifiera kluster med komplexa former skiljer den från andra klusteringsmetoder som ofta kräver förinställda klusterantal eller kan ha svårt att hantera brus.

2.3.4 Silhuettvärden

En metod som används för att jämföra och utvärdera kvaliteten på olika klustringsresultat är den så kallade silhuettmetoden [33]. Denna metod bedömer klustren baserat på två aspekter: sammanhållning, vilket mäter hur starkt ett element är kopplat till sitt eget kluster och separation, som mäter hur starkt ett element är kopplat till andra kluster. För varje objekt beräknas ett silhuettvärde som indikerar hur väl det passar in i sitt kluster. För att få ett övergripande värde på hur bra en klustring är kan ett genomsnitt över alla silhuettvärden tas.

Antag att datamängden består av n element som delas in i k olika kluster. För ett givet element $1 \leq i \leq n$ som tillhör kluster C_1 beräknas följande värden:

- Genomsnittlig separation från i till övriga element i C_1 :

$$a(i) = \frac{1}{|C_1| - 1} \sum_{j \in C_1 \setminus \{i\}} d(i, j)$$

- Genomsnittlig separation från i till alla element i ett kluster C_m , $m \neq 1$:

$$d(i, C_m) = \frac{1}{|C_m|} \sum_{j \in C_m} d(i, j)$$

- Det närmsta klustret till i :

$$b(i) = \min_{C_m, 1 \leq m \leq k} d(i, C_m)$$

- Silhuettvärdet för objekt i :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Här är $d(i, j)$ är ett avståndsmått, exempelvis det euklidiska avståndet. Om ett kluster C_m endast innehåller ett element i ansätter vi istället att $a(i) = 0$. För varje i kommer $s(i) \in [-1, 1]$, där ett högre värde innebär att i väldigt sannolikt ska tillhöra sitt kluster, då separationen inom klustret i detta fall går mot noll. Ett genomsnittlig silhuettvärde för hela klustringen kan tas fram med uttrycket

$$S = \frac{1}{n} \sum_{i=1}^n s(i). \quad (2.5)$$

För att jämföra olika klustringar kan klustringarnas genomsnittliga silhuettvärde beräknas, där den klustringen som har ett högre silhuettvärde anses vara bättre. Oftast tillämpas detta för att bestämma det optimala antalet kluster för k -means klustring, men det kan på samma sätt användas för andra parametrar och modeller för att hitta den konfiguration som ger bäst kluster [34].

Det är dock viktigt att betona att silhuettvärden inte är ett generellt mått på vad som innefattar ett bra kluster. Silhuettvärden är inte lämpliga för att jämföra klustringar som tas fram med dimensionsreducering mot utan dimensionsreducering då formeln för silhuettvärden direkt beror av datamängdens dimension [35]. En jämförelse av silhuettvärden riskerar därmed att ge missvisande resultat.

2.3.5 Scikit-learn

Scikit-learn är ett open-source-maskininlärningsbibliotek för programmeringsspråket Python [36]. Biblioteket är ett tillägg till SciPy och erbjuder en omfattande uppsättning verktyg och funktioner som underlättar för utvecklare att implementera olika maskininlärningsmodeller.

En av de framträdande funktionerna i scikit-learn är dess implementering av en mängd olika maskininlärningsalgoritmer, däribland oövervakad maskininläring. Dessa inkluderar algoritmer för klassificering, klustring och dimensionsreducering. Varje algoritm är utformad för att vara lätt att använda och optimerad för prestanda, vilket gör det möjligt för användare att experimentera med olika tillvägagångssätt.

Förutom maskininlärningsalgoritmerna erbjuder scikit-learn också omfattande stöd för dataförberedelse och modellvalidering. Det inkluderar verktyg för att hantera och förbereda data, såsom funktioner för skalning, normalisering och dimensionsreducering. Dessutom erbjuder biblioteket metoder för att utvärdera och validera modeller, vilket gör det möjligt för användare att bedöma prestanda och generaliseringsförmåga hos sina maskininlärningsmodeller.

3

Metod

I detta kapitel beskrivs de modeller och algoritmer som har använts för att utveckla den slutliga modellen. Kapitlet är uppdelat i de olika faserna som projektet genomgick för att konstruera den färdiga modellen. Varje delmoment redogör detaljerat för de specifika metoderna och tillvägagångssätten som tillämpades under projektets gång.

Metoden delades upp i tre olika delar: datahämtning, bildbehandling och maskininlärning. I avsnittet om datahämtning beskrivs processen för att skapa och samla in den använda datamängden. I bildbehandlingsavsnittet redogörs hur bilderna har behandlats för att minimera irrelevanta faktorer. I maskininlärningsavsnittet beskrivs implementationen av attributextrahering, dimensionsreducering och klustringsalgoritmer. Samtlig kod, inklusive testfilen, finns tillgänglig i github länkad i Appendix C.

3.1 Datahämtning

För datahämtning ur ALMA arkivet användes Python-biblioteket ALminer [13]. Som beskrivet i 2.2.1 möjliggör ALminer filtrering av data från ALMA-arkivet. På så sätt har riktad och relevant data extraherats som grund för fortsatt användning. I praktiken innebar detta en noggrann granskning av relevanta nyckelord innan dataextrahering genomfördes med syftet att minimera uteslutning av relevant data.

3.1.1 Filtrering

För att extrahera relevanta filer ur ALMA-arkivet innehållande protoplanetära skivor användes följande ALMA-nyckelord:

- “*disks around low-mass stars*”
- “*disks around high-mass stars*”

Från dessa valdes FITS-filer med namn innehållande:

- “*pbcor*”
- “*_sci*”
- “*cont*”

Dessa parametrar valdes för att få två-dimensionella vetenskapliga observationer av protoplanetära skivor. Vidare förklaras parametrarna under 2.2.3.

3.2 Bildbehandling

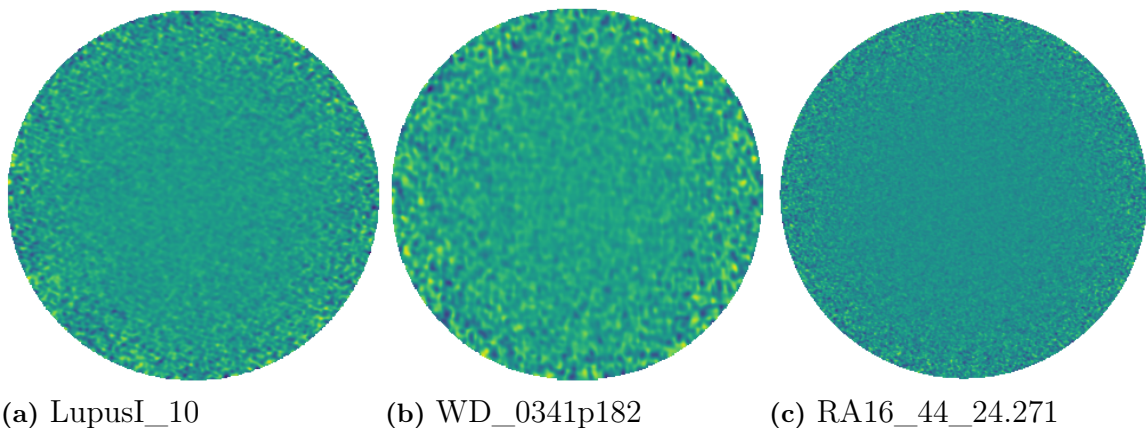
Innan bilderna klustrades med hjälp av maskininlärning implementerades en bildbehandling. Detta genomfördes för att minimera inverkan av faktorer som inte borde påverka klustringen. Några exempel på sådana faktorer är objektets position i bilden, bildens upplösning, brus och objektets inklination i himmelsplanet.

3.2.1 Beskärning av bild

Lokaliseringen av objekt och behandlingen av bilderna baserades på antagandet att det starkaste mätvärdet tillhörde objektet i bilden. På grund av det starkt kontrasterande bruset som finns längs kanterna på filer med “.*pbcor*”, riskerade en stark brus-pixel identifieras som objektet i bilden. För att undvika detta utfördes inledningsvis en beskärning av bilderna till en centrerad cirkel med halva diametern av den ursprungliga bilden. Detta eliminerade bruset samtidigt som det bevarade det intressanta objektet i bilden.

3.2.2 Selektiv bildeliminering

På en betydande andel av de hämtade bilderna från ALMA-arkivet syntes inga tydliga objekt. Några exempel på detta ses i figur 3.1.



Figur 3.1: Exempel på bilder utan tydligt objekt från ALMA-arkivet. Alla bilderna har filnamnstillägg *_sci*, *.cont* och *.pbcor* enligt 3.1.1. Objektet som bilden skall innehålla ses i bildtexten. Länkar till observationerna återfinns längst ner i textfilen i githuben, se Appendix C.

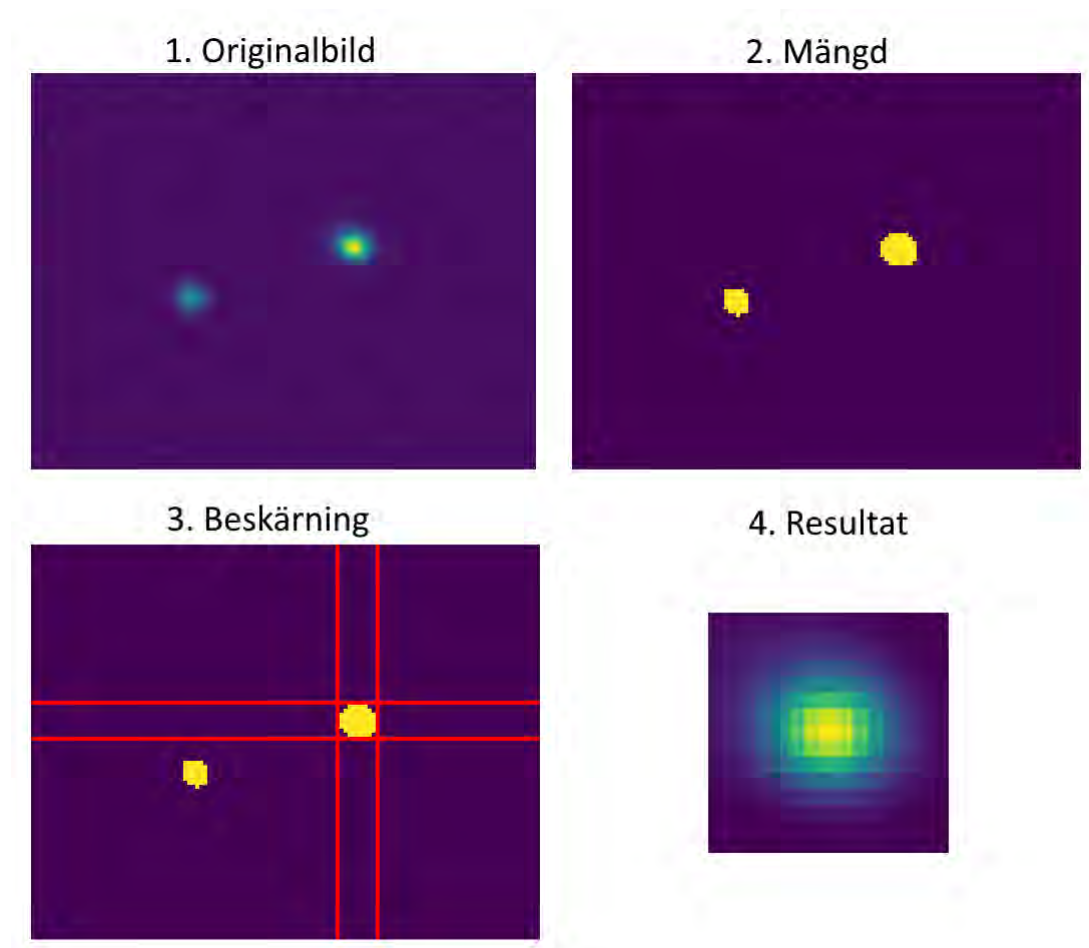
Innan resterande bildbehandling genomfördes filtrerades dessa bilderna bort. För att göra detta beräknades medelvärdet μ och standardavvikelsen σ av mätvärdena. Sedan eliminerades alla bilder som inte innehöll några pixlar med värden över $\mu + 7\sigma$. Valet av 7σ baserades på undersökning av bilderna som avlägsnades.

3.2.3 Förstoring

För att säkerställa att objektets storlek och position i bilden samt de varierande nivåerna av brus i bakgrunden inte skulle påverka klustringen, genomfördes en beskärning av bilderna till ett område som omringade objektet. Inledningsvis lokaliserades den protoplanetära skivan med antagandet att den ljusaste pixeln i den beskärda bilden tillhörde skivan. Förstoringsfunktionen såg till mängden M av pixlar P där

$$P_{val} \geq \max_{P \in bild} (P_{val}) \cdot (\mu + n \cdot \sigma)$$

för en varierbar faktor n , där μ och σ var anpassade enligt ovan. Bilden beskars till närmaste rader respektive kolumner på vardera sida om maximumpunkten som inte innehöll pixlar i M , se figur 3.2. Detta steg upprepades tills bilden inte längre minskade i storlek. Runt området sparades även en ram med tjocklek proportionell mot bildens storlek i både x - och y -led. För de slutgiltiga testerna användes 20% av bildens storlek till tjocklek av ramen.



Figur 3.2: Illustration av förstoringprocessen där resultatbilden inkluderar en ram.

Olika val av n undersöktes för att hitta det som gav bäst slutresultat i klustringen. Ett mindre n ökade risken för att inkludera brus runt skivorna, medan ett större n riskerade att delar av skivorna beskars bort.

Flertalet enklare varianter av förstoringfunktioner undersöktes. En variant var förstoring till den minsta rektangel där hela M inkluderades. Denna metod fungerade inte tillfredsställande om bilden innehöll två objekt likt 3.2. En annan variant var förstoring med utgångspunkt i maximumpunkten där den första pixel i respektive riktning som inte ingick i M hittades. Bilden beskars sedan till rektangeln mellan de fyra pixlarna. Denna metod fick istället problem med att ringar runt stjärnorna skars bort. Den slutgiltiga varianten var en vidareutveckling av dessa två metoder.

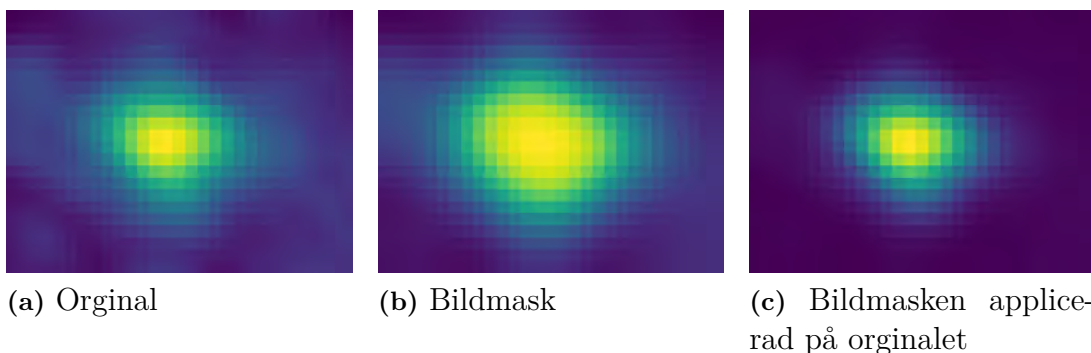
3.2.4 Maskering med oskärpa

För att ytterligare framhäva objektets form i bilderna och skapa en tydligare kontrast till bakgrunden implementerades en oskarp maskering. Detta genomfördes genom att skapa en kopia av den förstörade bilden som erhöles efter utförandet av 3.2.3, vilken fungerade som bildmask. I bildmasken normaliserades alla mätvärden längs en linjär skala mellan noll och ett justerbart maximalt värde. Därefter justerades bildmaskens värden enligt

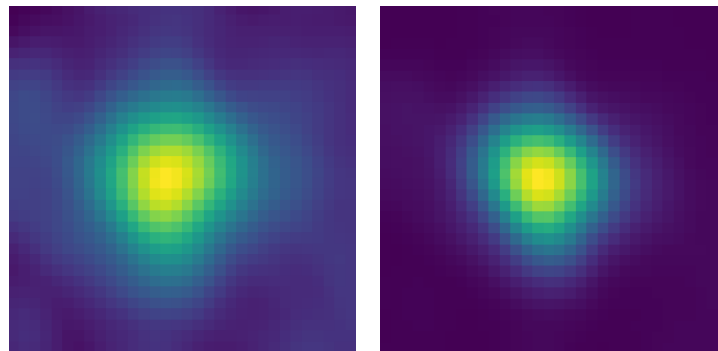
$$P_{val} = \begin{cases} P_{val}, & P_{val} \leq 1 \\ 1, & P_{val} > 1 \end{cases}$$

Detta resulterade i en avgränsning när bildmasken applicerades, där de högre värdena, som med stor sannolikhet tillhörde objektet, påverkades mindre än de lägre värdena. De nya värdena i bildmasken upphöjdes sedan till en exponent för att skapa en mer markant avtagande effekt.

Därefter applicerades ett gaussiskt filter för att göra bildmasken suddig och för att motverka skapandet av alltför skarpa kanter som inte skulle vara representativa för den ursprungliga formen. Slutligen multiplicerades bildmasken med den förstörade bilden, vilket resulterade i en ny bild som liknar originalet, dock med mindre kringliggande bakgrundsbrus. Denna process illustreras i figur 3.3 och figur 3.4.



Figur 3.3: Exempel på hur bildmaskeringen påverkar den förstörade bilden av objektet TWA_7.



(a) Slutlig bild utan genomförandet av bildmaskering
 (b) Slutlig bild med genomförandet av bildmaskering

Figur 3.4: Exempel på hur bildmaskeringen påverkar den slutliga bilden av objektet TWA_7.

3.2.5 Skaländring och rotation

Problemet med att de förstörade bilderna fick olika upplösning löstes genom en skaländring. För att bestämma värdet på varje ny pixel interpolerades mellan värdena på de fyra närliggande pixlarna i den gamla bilden. För en pixels position användes dess centrum. En önskvärd viktning var normaliserad och större för närliggande pixlar, med värde 1 om positionen var densamma. För denna viktning användes en “invers distans”, för vilken flera olika varianter undersöktes. Några varianter för vikten v av nedre vänstra pixeln är

$$v = \frac{1}{s} \left((1-x)^2 + (1-y)^2 \right) \quad (3.1)$$

$$v = \frac{1}{s} \left(\frac{1}{x^2 + y^2} \right) \quad (3.2)$$

$$v = (1-x)(1-y) \quad (3.3)$$

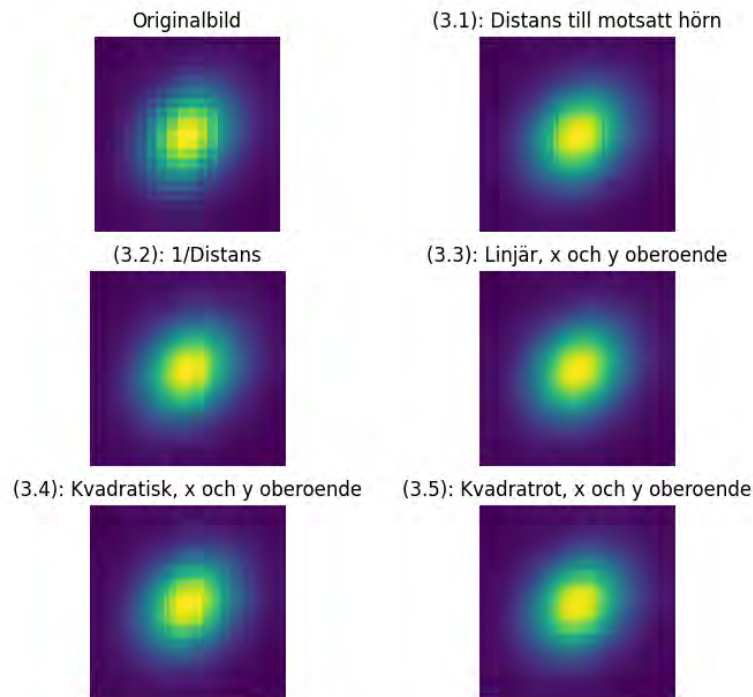
$$v = \frac{1}{s} \left((1-x)^2 (1-y)^2 \right) \quad (3.4)$$

$$v = \frac{1}{s} \left(\sqrt{(1-x)(1-y)} \right), \quad (3.5)$$

där övriga vikter erhålls genom ett variabelbyte. Här är x och y avståndet i antal pixlar i x - respektive y -led och s summan av vikterna. Pixelvärdet i den ändrade bilden, $P_{ny,val}$ gavs sedan av

$$P_{ny,val} = \sum_P v_P P_{val},$$

där det summeras över de fyra närliggande pixlarna. En jämförelse av de olika varianterna av invers distans ses i figur 3.5.



Figur 3.5: Skaländring med olika varianter av invers distans.

Variant (3.3) gav bäst visuella resultat vid prövning av flertalet bilder och används framöver.

Ytterligare ett problem som hanterades av skaländringen var skivornas inklinering i himmelsplanet. En disk med hög inklinering uppstod som en oval och riskerades att grupperas baserat på inklinationen. Inklinering som innebar att bilden inte blev kvadratisk efter förstoringen hanterades direkt av skaländringen. Utöver detta behövde endast disken “dras ut” längs med diagonalerna, vilket kan ses som att den roterades över bildens diagonaler. Rotationen utfördes genom att manipulera var i den ursprungliga bilden pixlarna hämtade sina värden under skaländringen.

Inledningsvis hittades en rotationsfaktor k_R som berättade hur stor inklinering skivan hade och kring vilken diagonal den behövde roteras. Låt l_1 vara längden av diagonal 1, d_1 och l_2 längden av diagonal 2, d_2 . Då ges rotationsfaktorn av

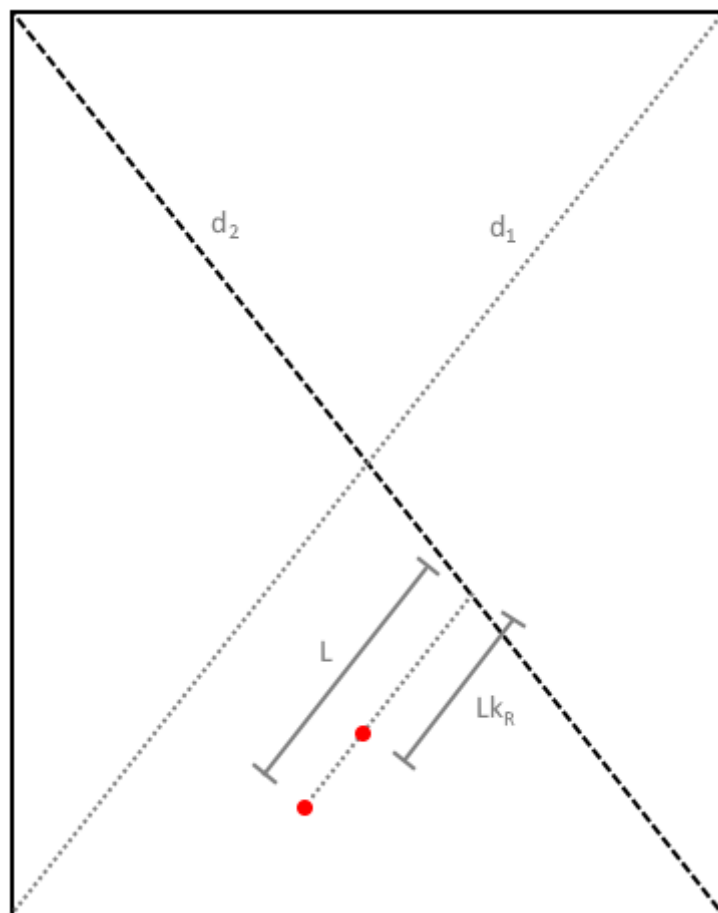
$$k_R = \frac{l_1}{l_2}.$$

Längderna av diagonalerna hittades genom att genomföra en skaländring utan rotation som ovan, där endast värden för pixlar längs diagonalerna beräknades. Med utgångspunkt i hörnen hittades det första motstående par av pixlar P_i och P_{-i} längs diagonalen sådana att

$$P_{i,val} + P_{-i,val} \geq 2 \cdot \max_{P \in B}(P_{val}) \cdot (\mu + 7\sigma),$$

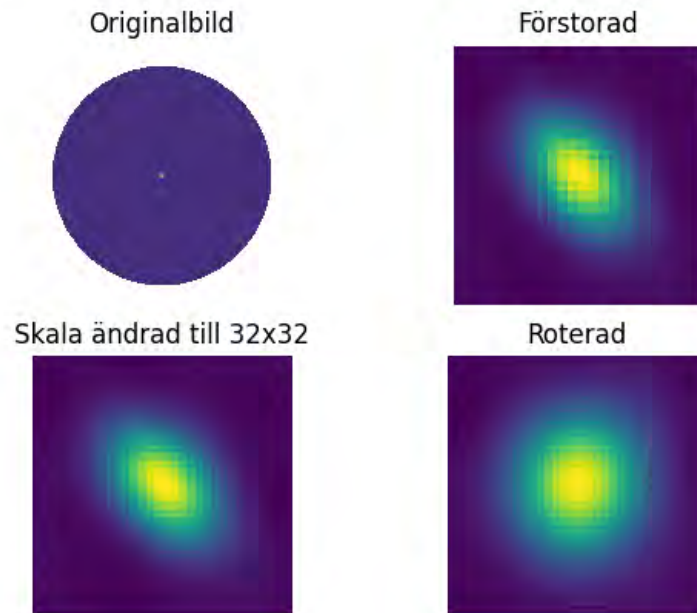
där k var en varierbar faktor likt under 3.2.3 samt B refererar till bilden före skaländring. Längden av diagonalen beräknades som antalet pixlar från P_i till och med P_{-i} .

Rotationsfaktorn användes sedan för att skifta punkten i den gamla bilden, varifrån pixlarna plockade sina värden, närmare diagonalen runt vilken skivan skulle roteras. Detta innebar att skivan “drogs ut” tills den blev cirkulär i bilden. För rotation kring diagonal 1, där $k_R > 1$, flyttades punkten i riktning parallell med d_2 tills avståndet från d_1 var $\frac{1}{k_R}$ av initialavståndet. På liknande sätt genomfördes rotation kring d_2 , se figur 3.6. Efter att punkten justerats beräknades värdet efter skaländringen precis som tidigare, genom att interpolera mellan pixlarna närmast den nya punkten.



Figur 3.6: Illustration av hur positionen där värdet i bilden hämtas och justeras vid rotation kring d_2 .

Metoden för att rotera bilderna baserades på att skivorna var cirkulära och centrerade i bilden. Vid korrekt förstoring borde däremot en cirkulär skiva vara centrerad. Syftet med metoden var att på ett enkelt sätt ta bort delar av inklanationens påverkan på klustringen. För att kontrollera rotationens påverkan på resultatet sparades



Figur 3.7: Exempel på förstoring, skaländring och rotation av en skiva, samt skaländring utan rotation.

k_R för varje bild. De slutgiltiga klustrena med och utan rotationen jämfördes sedan baserat på spridningen av bilder med ovanligt stort eller litet k_R .

3.2.6 Sparande av bilddata

För att möjliggöra en mer nyanserad undersökning av klustringen sparades information om bilderna i en json-fil ¹. Informationen som sparades var namnet på objektet, bildens storlek efter förstoringen samt rotationsfaktorn k_R . Namn på objektet extraherades från nyckelord i FITS-filerna och användes för att undersöka om bilder på samma objekt hamnade i samma kluster. Detta kontrollerades sedan via en testfil som skrev ut vilka kluster samma objekt tillhörde, ett exempel av detta kan ses i Appedix D. Övrig information sparades för att möjliggöra undersökning av skaländringen och rotationens påverkan på slutresultatet.

3.2.7 Omvandling till JPG

Innan användning i maskininlärningsmodellen omvandlades bilderna från matriser till JPG. Maskininlärningsmodellen läste in data som matriser, så konverteringen av bildfilerna till JPG var inte nödvändigt ur ett maskinlärnings syfte, men gjordes för

¹JavaScript Object Notation (JSON) är ett textbaserat filformat som används för att lagra och utbyta data.

att kunna visualisera de processerade bilderna. Varje pixel i bilden korresponderade mot ett matriselement, där högre värden fick en ljusare färg.

3.3 Maskininlärning

Efter bildbehandlingen implementerades maskininlärningen. Processen innefattade attributextrahering, dimensionsreducering och klustring.

3.3.1 Extrahering av bildattribut

De fyra attributextraheringsmetoderna som beskrevs under avsnitt 2.3.1 implementerades som funktioner i Python. För attributsextraktion med CNN användes VGG16, ett färdigtränat faltningsnätverk från paketet Keras [37], där de yttersta lagret eliminerats. I funktionen behandlades även varje given bild med en tillhörande bildbehandlingsfunktion från Keras.

Implementationen av attributsextraktion med pixelvärden och DCT var liknande. Pixelattributsberäkningen utfördes genom inläsning av en bild som en matris, vilken sedan plattades ut. För DCT-attributen genererades en matris av DCT-transformerade pixelvärden beräknade med en funktion från paketet SciPy [38].

För extraktion med HOG användes en färdig funktion från paketet Scikit-Image [39], där bilden delades in i block om 8×8 pixlar och endast en cell per block användes.

3.3.2 Dimensionsreducering

De två beskrivna dimensionsreduceringsmodellerna implementerades i ett separat kodblock för att göra det möjligt att köra resten av programmet utan dimensionsreducering. Både PCA och t-SNE implementerades med funktioner från paketet scikit-learn [36].

3.3.3 Klustring

Efter att attributextrahering och dimensionsreduceringen genomförts byggdes en oövervakad maskininlärningsmodell upp, där k -means, GMM och DBSCAN implementerades med hjälp av scikit-learn biblioteket som beskrivs i avsnitt 2.3.5.

K -means genomfördes med "*n_clusters*" inställt som ett givet värde. Därefter beräknades klustringen och etiketter tilldelades varje kluster. En liknande metod användes för GMM, men istället för att ett givet värde på antalet kluster användes Bayesian Information Criterion (BIC), utvecklat av Schwarz [40], för att automatisera denna process. Det kluster med lägst poäng tilldelades som "*n_cluster*" i klusterberäkningen.

Därefter utvecklades DBSCAN som en ytterligare metod för klustring. DBSCAN

använde inte ett fördefinierat antal kluster, vilket beskrivs i avsnitt 2.3.3.3, utan olika värden på parametrarna ε och “*min_samples*” testades. Algoritmen skapade sedan varierande antal kluster beroende på vad parametrarna sattes till.

3.4 Utvärdering

För att kunna utvärdera de olika modellerna som hade implementerats söktes ett konkret sätt att jämföra de klustringar som producerats. För att jämföra attributextraheringsmetoder beräknades silhuettvärden med k -means och utan dimensionsreducering, där ett högre värde indikerade att attributextraheringsmetoden var bättre. En subjektiv bedömning gjordes även genom att klustringsresultaten analyserades för de olika modellerna. Den subjektiva bedömningen byggde bland annat på storleken på klustrena, storlekens varians, om bilderna i samma kluster liknade varandra och om det förekom liknande bilder i andra kluster.

Jämförelse av klustringsmetoder gjordes till största del subjektivt, på samma sätt som för jämförelsen av attributextraheringsmetoder samt genom att analysera om olika observationer av samma objekt placerades i samma kluster. För att göra jämförelsen enklare skapades en testfil med en funktion som beräknade om samma objekt fanns i flera bilder och vilka kluster som detta objekt i så fall tillhörde. Dessutom skrev funktionen ut hur varje objekt hade roterats och förstörats för varje kluster objektet hamnat i. Detta utgjorde en grund för att bestämma hur rimliga klustrena var för de olika metoderna och parametrarna.

4

Resultat

I detta kapitel presenteras resultatet av de tester som gjorts. De implementerade maskininlärningsmetoderna k -means och DBSCAN används för att skapa kluster med olika typer av datamängder. Bilderna i dessa olika datamängder genomgår flera olika varianter av bildprocessering. Därefter testas resultaten av olika kluster med hjälp av ett program för att avgöra vilka objekt de olika klustrena består av. Målet med testerna är att se hur olika parametrar både i bildprocesseringen och maskininlärningsalgoritmerna påverkar de returnerade klustrena.

4.1 Datahämtning och brusfiltrering

Med de inställningar som beskrivs i metoden [3.1.1](#) så resulterade ALminer sökningen i 5746 observationer av 3749 unika objekt. Av dessa så laddades 959 stycken ner varav 360 ansågs ha för mycket brus för att kunna nyttjas enligt kapitel [3.2.2](#). Detta resulterade i 599 bilder som användes för den slutgiltiga behandlingen och klustringen. Samtliga observationer som används i projektet återges i projektets GitHub.

4.2 Bildprocessering

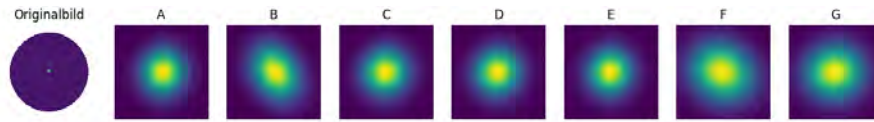
För att undersöka bildbehandlingen skapas flera olika datamängder. Varje datamängd skapas utifrån flera olika parametrar varav vissa är förutbestämda och vissa skiljer sig mellan datamängderna. De parametrar som skiljer sig mellan datamängderna är faktorn k , huruvida bilderna använder sig av en oskarp maskering och huruvida bilderna roteras eller endast skaländras.

Faktorn k bestämmer hur mycket bilden zoomas in, vilket beskrivs i kapitel [3.2.3](#). Parametern maskering behandlar huruvida ett suddigt maskerande filter, som minskar påverkan av brus kring målen implementeras eller inte, vilket återges i metodkapitlet [3.2.4](#). I metoden [3.2.5](#) beskrivs att samtliga bilder skaländras, men den sista parametern beskriver ifall bilden, utöver skaländring, också roteras så som beskrivits i samma kapitel. I följande tabell [4.1](#) listas parametrarna för de olika datamängderna, namngivna med A-G, som används i analysen.

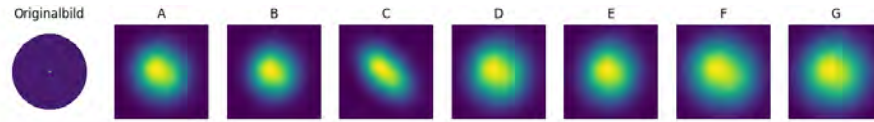
Datamängd	n	Maskering	Roterad
A	3	Nej	Ja
B	3	Ja	Ja
C	5	Nej	Nej
D	5	Nej	Ja
E	5	Ja	Ja
F	7	Nej	Ja
G	7	Ja	Ja

Tabell 4.1: De olika datamängderna som används med deras respektive parametervärden n , maskering och roterad.

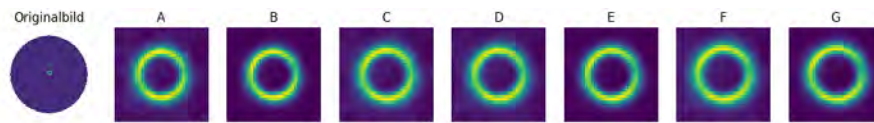
I figur 4.1 nedan följer ett urval av filer processade enligt vardera av datamängderna. Filerna är utvalda för att illustrera hur bildbehandlingen behandlar olika typer av protoplanetära skivor och olika mängd brus. Datamängderna är markerade med namn ovanför de processerade bilderna. Namnen är tagna från den ovanstående tabellen 4.1. Bilden längst till vänster återger originalbilden direkt från ALMA-arkivet och de resterande bilderna är originalbilden behandlade utifrån den namngivna datamängden.



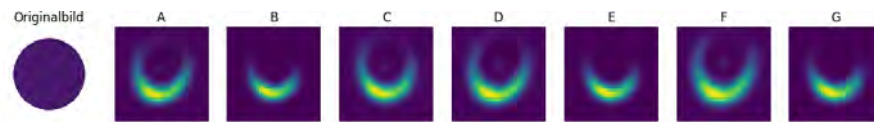
(a) En skiva där ingen rotation krävs. Objekt: RA16_28_16.498



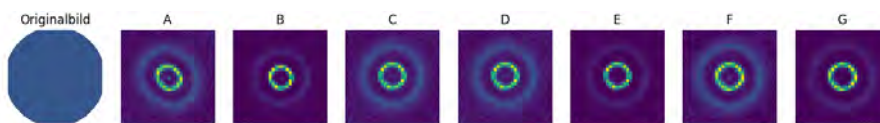
(b) Skiva i vinkel som roteras. Objekt: HD_34282



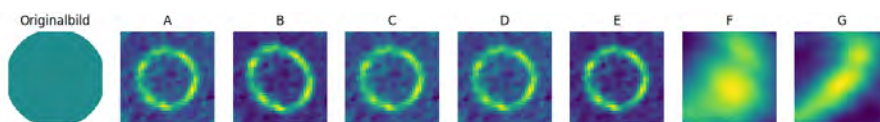
(c) Skarp ring. Objekt: 2MASS_J16042165-2130284



(d) Ring med svagare sida. Objekt: HD_142527



(e) Objekt med yttre ring. Objekt: HD_169142



(f) Bild med mycket brus. Objekt: HD_169142

Figur 4.1: Resultat av bildbehandlingen för olika ursprungsbilder, där bilden längst till vänster är originalbilden och de resterande är bilder behandlade i enlighet med dess korresponderade datamängd återgiven ovanför bilden.

4.2.1 Rotation

I syfte att undersöka inklanationen i himmelsplanets påverkan på klustringen, följer här en presentation av kluster som innehåller flest av de roterade stjärnorna. I datamängd D har 86 av 599 bilder en rotationsfaktor $k_R \geq \frac{3}{2}$ eller $k_R \leq \frac{2}{3}$, se metod 3.2.5. Låt M vara mängden av dessa bilder. I tabell 4.2 nedan visas antal kluster som innehåller bilder ur M , kallas antal kluster i tabellen, och antalet bilder i de två kluster som innehåller flest bilder ur M , kallas maxantal i tabellen. Detta visas för olika maskininlärningsmodeller och för datamängd C och D.

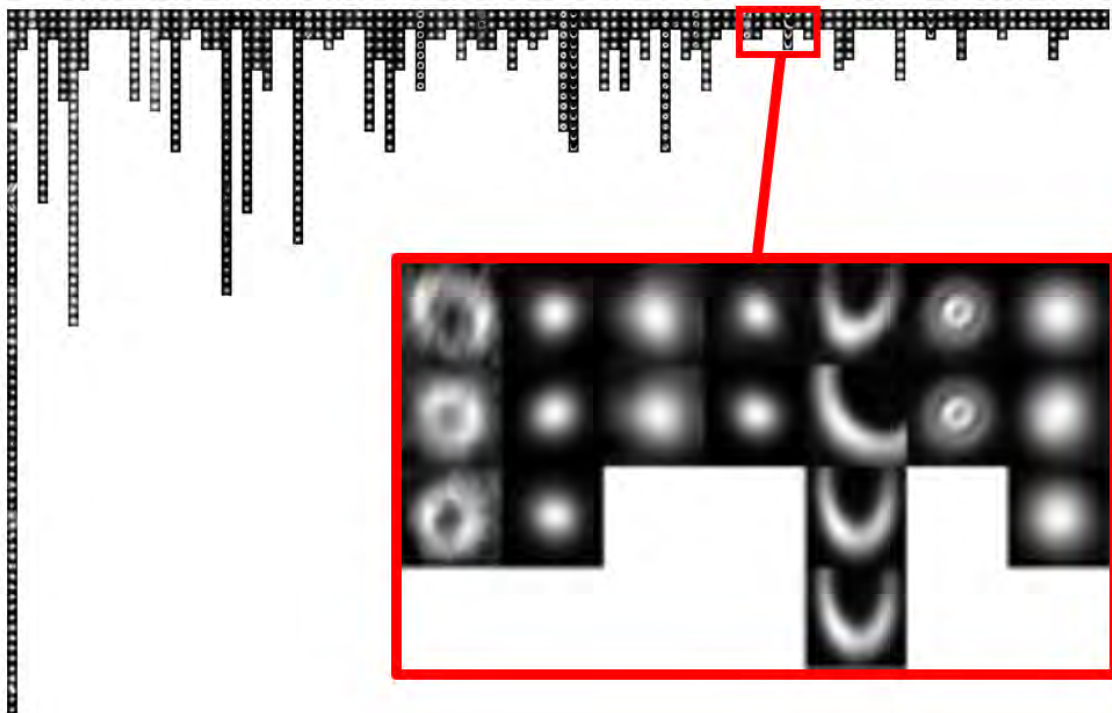
Modell	Datamängd	Rotation	Antal kluster	Maxantal
k -means, t-SNE, $k = 70$	C	Nej	24	12 10
	D	Ja	41	5 5
k -means, $k = 70$	C	Nej	26	13 9
	D	Ja	39	8 8
DBSCAN, t-SNE, $\varepsilon = 1$	C	Nej	26	13 8
	D	Ja	36	11 6

Tabell 4.2: Modell, datamängd, huruvida datamängden tillämpar rotation, antal kluster som innehåller bilder ur M och antalet bilder i de två kluster som innehåller flest bilder ur M , kallas maxantal i tabellen.

4.3 Maskininlärningsmodeller

Då vår modell beror av många olika parametrar kan ett stort antal klustringsresultat tas fram. I denna del presenteras endast de resultat och bilder som vi tycker visar tydligast resultat. Fler resultat återfinns i rapportens Appendix A.

Resultatbilderna över kluster består av kolumner av bilder, där varje kolumn motsvarar ett kluster. Ett exempel på detta illustreras i figur 4.2.

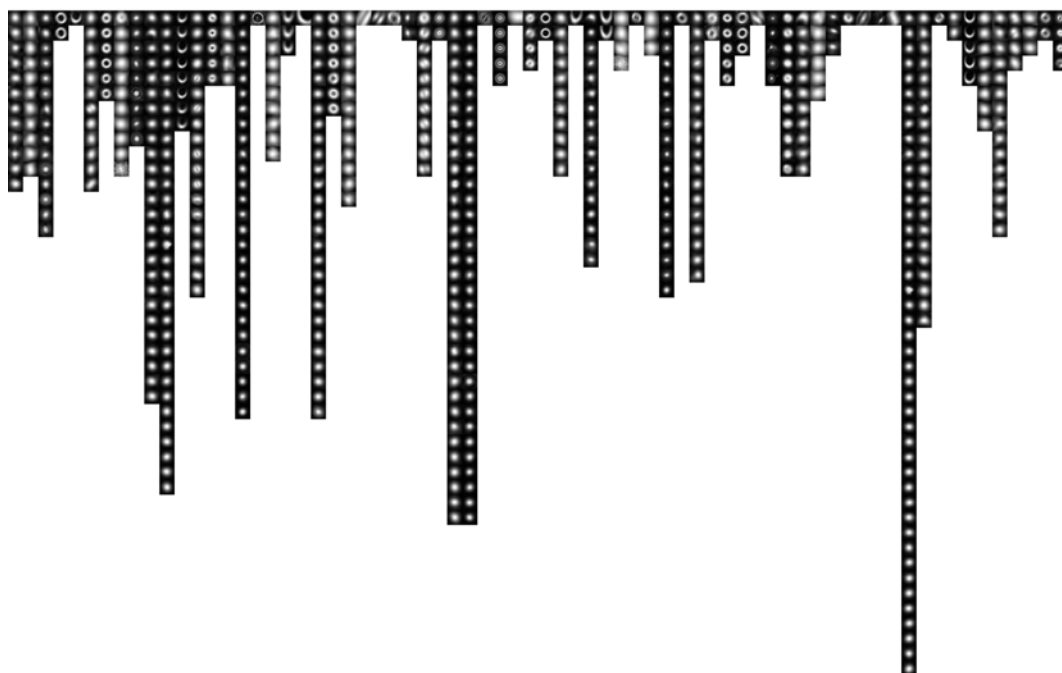


Figur 4.2: Exempel på strukturen av resultatbilderna. Den röda rutan visar en förstoring av några kolumner/kluster.

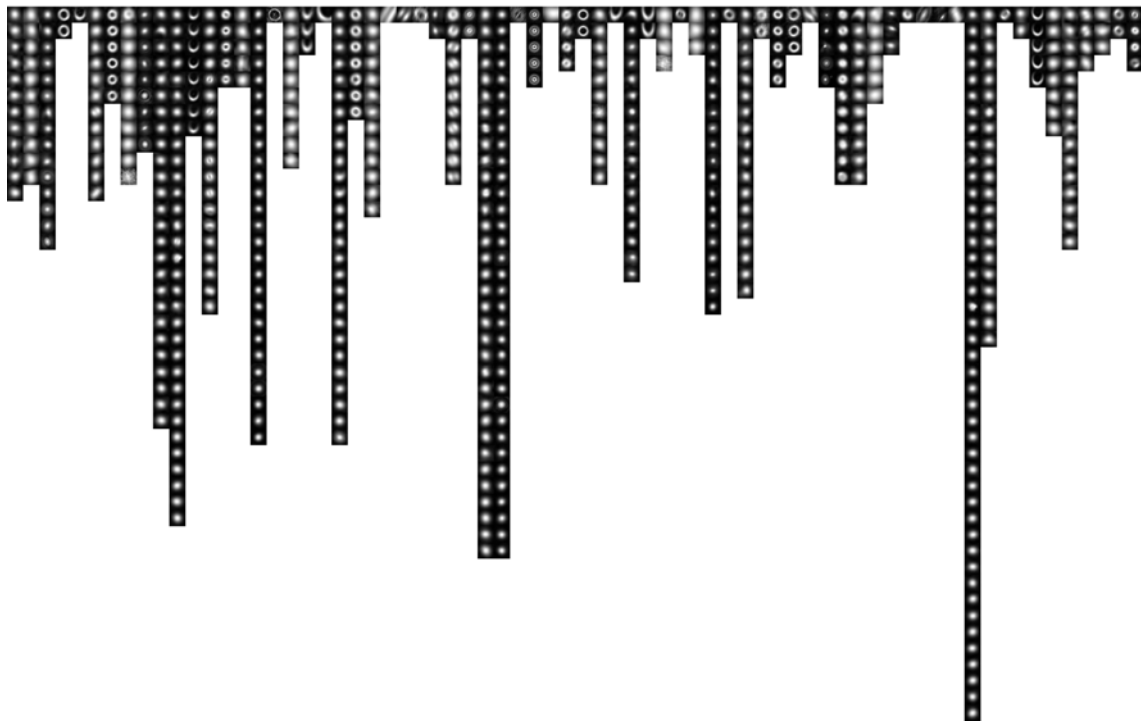
4.3.1 Jämförelse av attributextraheringsmodeller

Kluster har beräknats med k -Means, $k = 70$ och ingen dimensionsreducering på datamängd D . Nedan återfinns dessa figurer för de fyra olika attributextraheringsmodeller som utvecklats; CNN, pixelvärden, DCT och HOG. I figur 4.3 användes pixelvärdesattribut som attributextraheringsmodell, i figur 4.4 användes DCT, i figur 4.5 användes HOG och i figur 4.6 användes CNN.

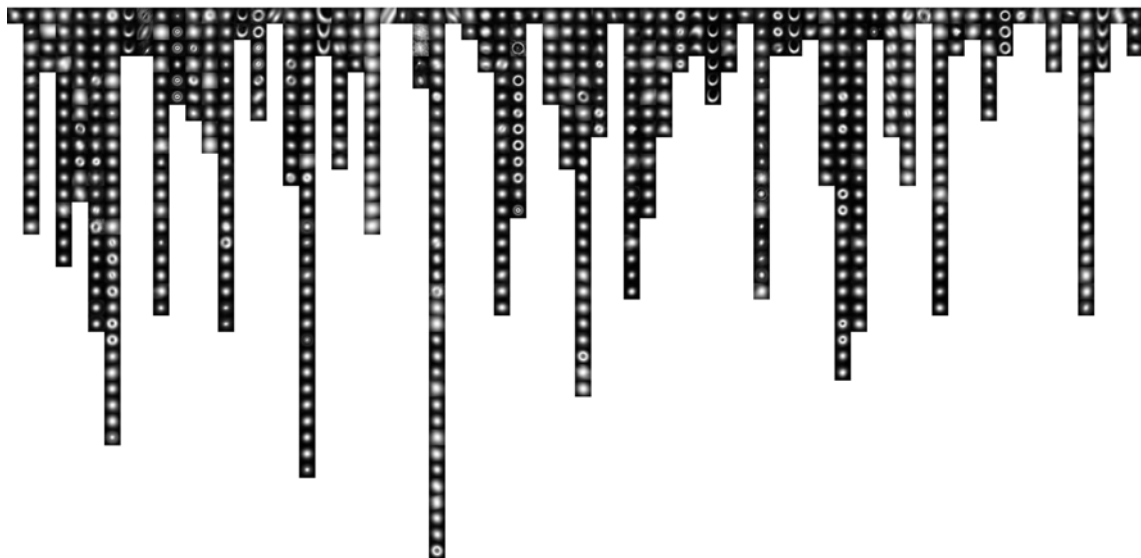
Slutligen beräknades silhuettvärden för samtliga kluster. Detta i syfte att undersöka vilka kluster som är rimligast. Mer detaljerad fakta om silhuettvärden kan hittas i Teori 2.3.4. Resultaten av detta redovisas i tabell 4.3 nedan, där attributextraheringsmodell deklarerar till vänster och det beräknade silhuettvärdet till höger.



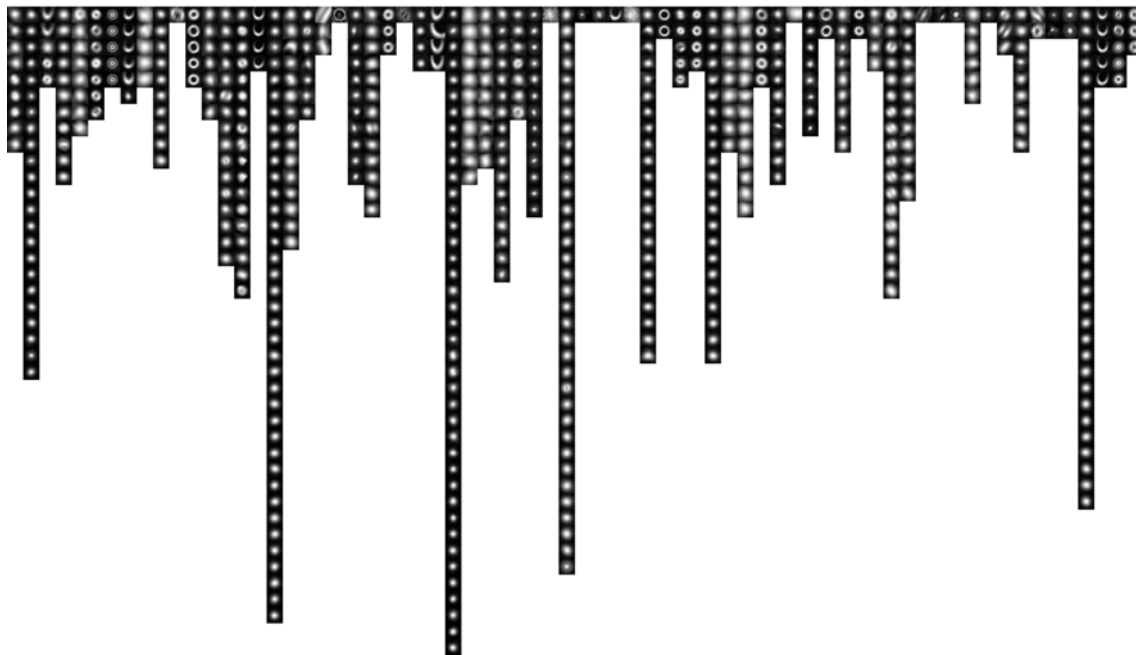
Figur 4.3: Kluster med pixelvärdesattribut.



Figur 4.4: Kluster med DCT-attribut.



Figur 4.5: Kluster med HOG-attribut.



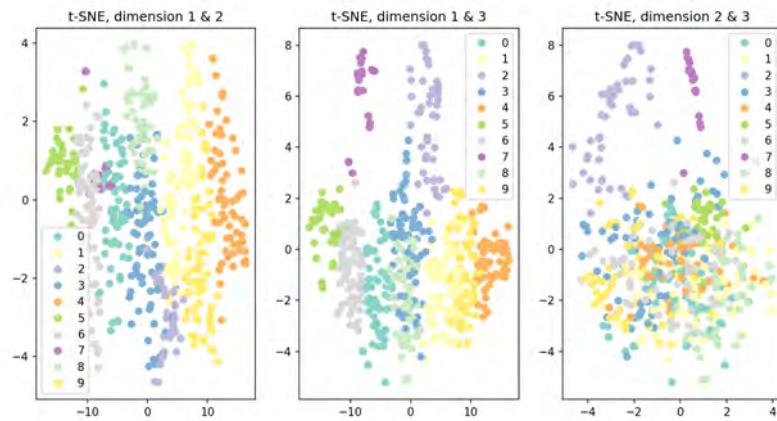
Figur 4.6: Kluster med CNN-attribut.

Pixelvärde	0.15345528677127543
DCT	0.15345528677127543
HOG	0.018601207734905444
CNN	0.084240004

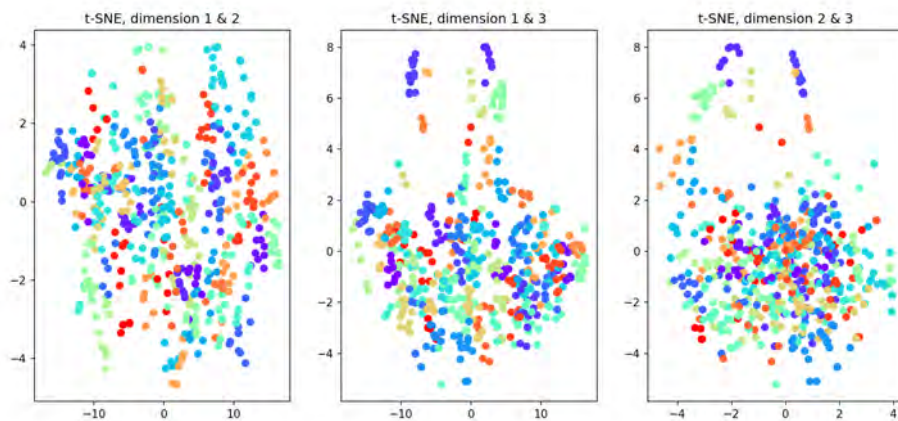
Tabell 4.3: Silhuettvärden av klustrerna för de olika attributextraheringsmetoderna.

4.3.2 Dimensionsreducering med t-SNE

Följande figurer 4.7 och 4.8 visar resultaten av dimensionsreducering av attributvektorn från DCT till 3 dimensioner med t-SNE. t-SNE har använts med parametrarna “*n_components*” = 3 , “*verbose*” = 1 och “*perplexity*” = 40. För att lättare visualisera den tredimensionella bilden presenteras den som tre separata figurer, där varje figur har två av de tre dimensionerna. Punkterna i figurerna representerar en bild och färgerna är representativa för det kluster som bilden blivit tilldelad efter att *k*-means kört. I figur 4.7 användes *k* = 10 för *k*-means och i figur 4.8 användes *k* = 70.



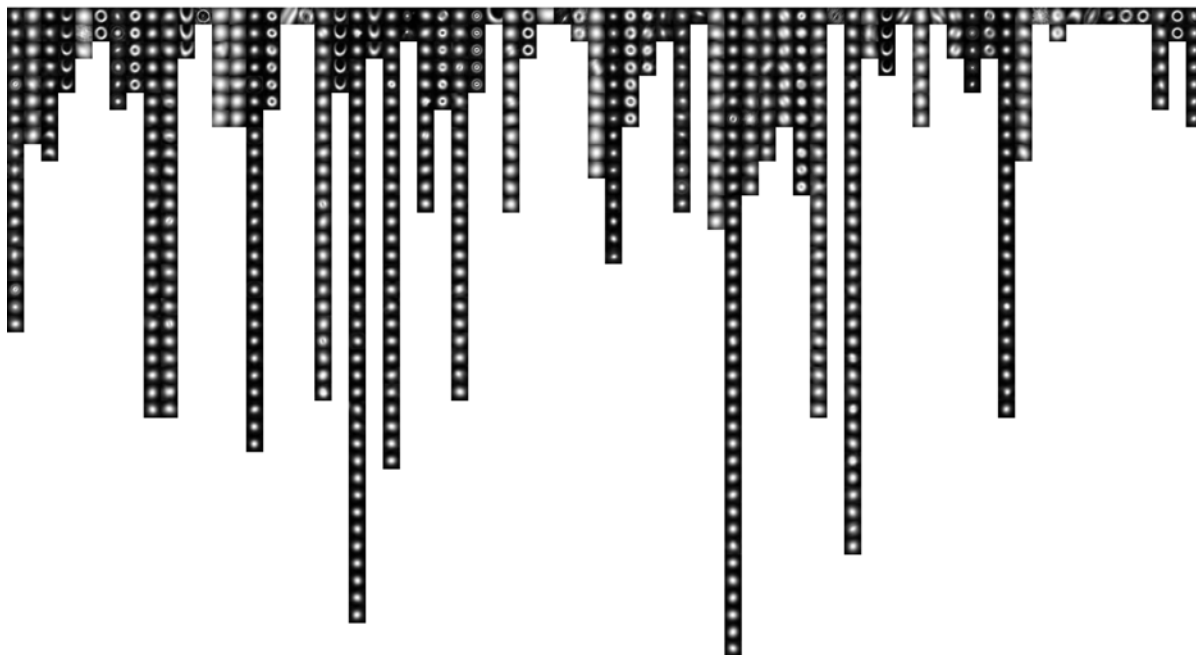
Figur 4.7: t-SNE för $k = 10$ kluster, där en prick representerar en bild och de olika färgerna representerar de olika klustererna.



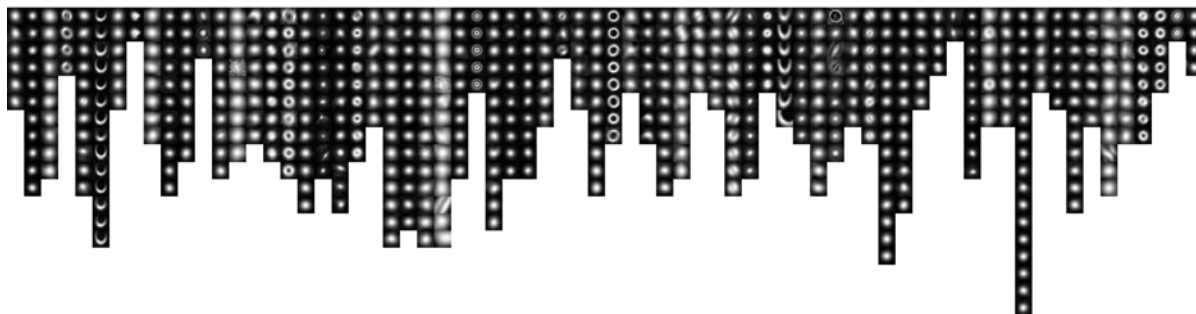
Figur 4.8: t-SNE för 70 kluster, där en prick representerar en bild och de olika färgerna representerar de olika klustererna.

4.3.3 Klusterallokering med k -means

Varje kolumn i figur 4.9 och 4.10 består av bilder placerade i samma kluster fördelade med maskininlärningsmetoden k -means, vilket beskrivs i 2.3.3.1. Skillnaden mellan de två figurerna är att i figur 4.10 klustras bilder som behandlats med t-SNE, beskrivet i avsnitt 2.3.2.1, och i figur 4.9 klustras bilder som inte har dimensionsreducerats.



Figur 4.9: Kluster på datamängd D med inställningar: k -means $k=70$ med DCT utan t-SNE.

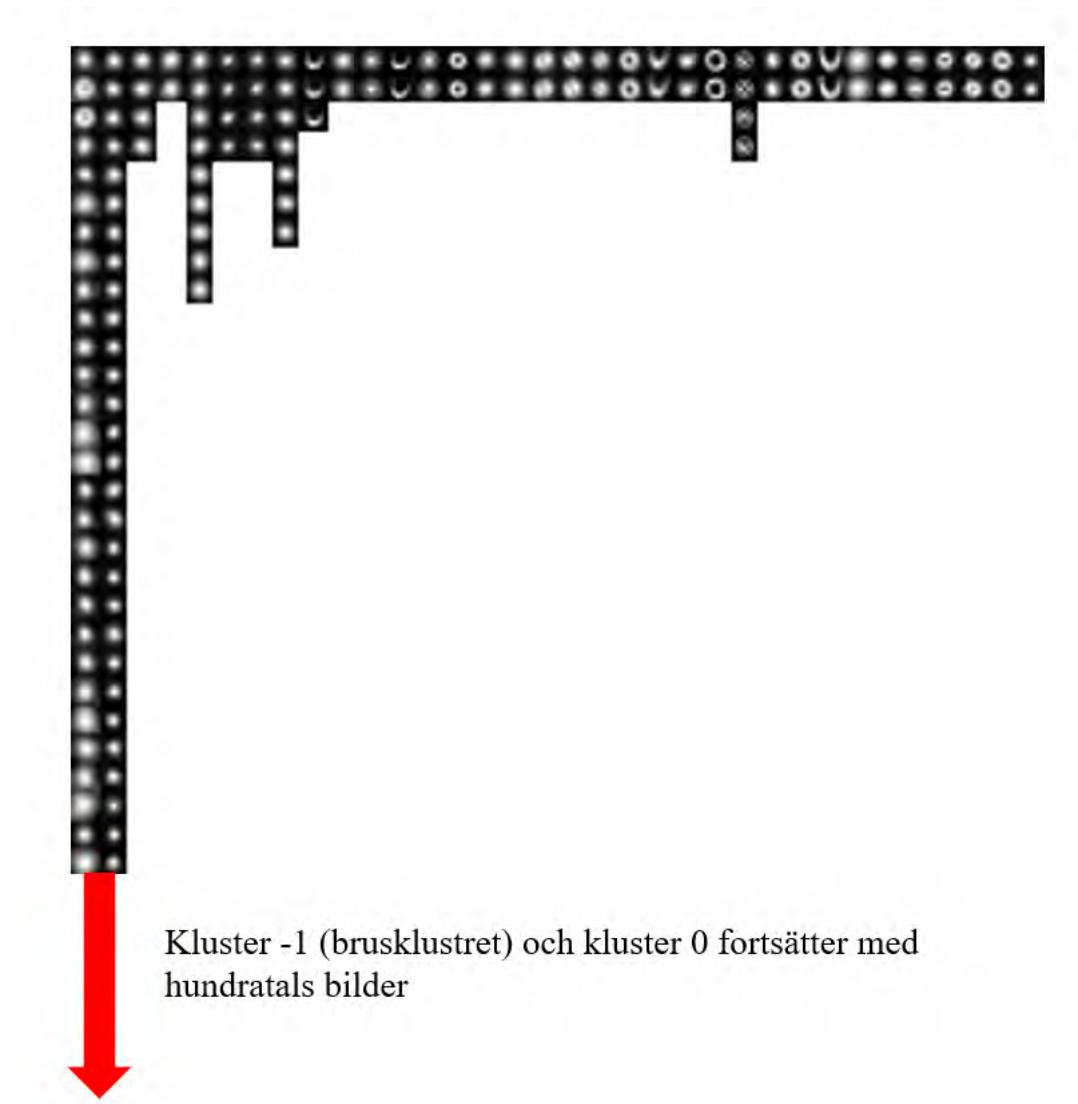


Figur 4.10: Kluster på datamängd D med inställningar: k -means $k=70$ med DCT och t-SNE

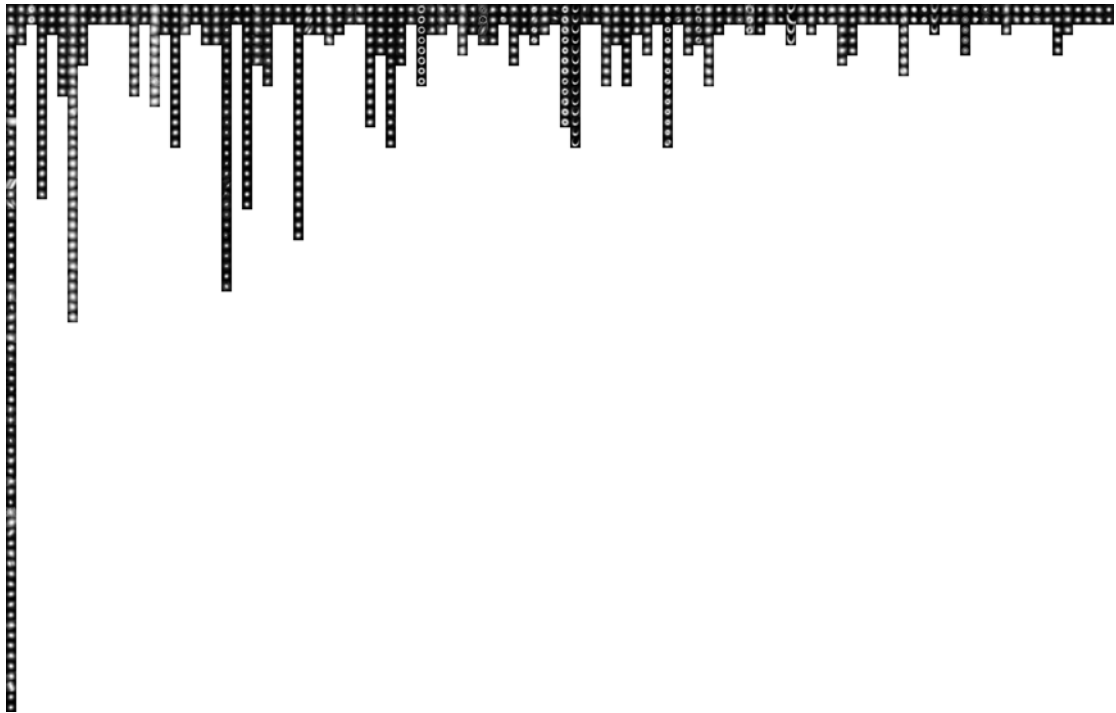
4.3.4 Klusterallokering med DBSCAN

Varje kolumn i figur 4.11 och 4.12 består av bilder placerade i samma kluster med hjälp av DBSCAN. I det första klustret -1 , kolumnen längst till vänster, hamnar alla bruspunkter.

Skillnaden mellan klustersamlingarna från DBSCAN är att figur 4.12 klustrar bilder vilka behandlats med t-SNE och figur 4.11 klustrar bilder som inte har dimensionsreducerats. Originalversionen av figur 4.11 är för stor för att kunna presenteras i resultatavsnittet och är därmed redigerad. Den röda pilen representerar var bilden har blivit beskuren.



Figur 4.11: Kluster på datamängd D med inställningar: DBSCAN $\varepsilon=1$ och $\text{min_samples}=2$ med DCT utan t-SNE.



Figur 4.12: Kluster på datamängd D med inställningar: DBSCAN $\varepsilon=1$ och $\text{min_samples}=2$ med DCT och t-SNE.

4.3.5 Utvärdering av tester

I Appendix B går det att observera hur objekt fördelats i olika kluster för varje metod när de förekommer mer än en gång i vår datamängd. Denna analys utgör en av grunderna för vår utvärdering av vilka kluster som presterar bäst, vilket förklaras i avsnitt 3.4.

5

Diskussion

I detta kapitel analyseras och diskuteras projektet utifrån dess metod, resultat, jämförelser med tidigare arbeten samt dess potential för vidareutveckling och de begränsningar som identifieras.

5.1 Datahämtning

Med inställningarna beskrivna i avsnitt 3.1.1 resulterade sökningen i 5746 observationer på 3749 unika objekt. Många av dessa observationer saknar “.pb”-filer vilket resulterar i att de inte kan användas i bildbehandlingen. Istället valdes användning av endast “.pbcor”-filerna vilket ger en större datamängd.

5.2 Bildprocessering

Det blev tidigt tydligt att när arkivbilder användes som inmatning till en maskininlärningsmodell utan någon förbehandling sorterades de endast baserat på ljusstyrkan av bruset i bilden. Bildprocesseringens primära syfte är därför att standardisera bilderna för att möjliggöra en effektiv klustring baserat på de protoplanetära skivornas form, istället för brus eller andra faktorer som storlek eller position.

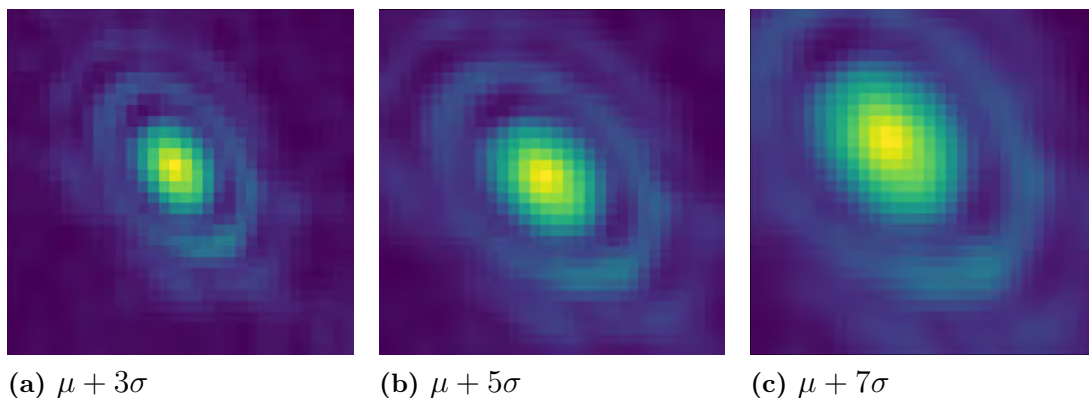
5.2.1 Beskärning

Eftersom bruset längs kanten av bilderna är starkt varierande beskärs bilderna till en radie hälften så stor som den ursprungliga bilden. Beskärningen eliminerar det starkt varierande bruset. En alternativ lösning som övervägdes var att använda observationernas “.pb”-filer för att erhålla likformigt brus över hela bilden. Likt tidigare nämnt fanns svårigheter i att hitta “.pb”-filer för alla “.pbcor”-filer. Istället valdes att beskära bilderna då det i praktiken endast påverkade kringliggande brus och inte eliminerade någon värdefull data.

I vissa specifika fall orsakar dock beskärningen problem. Om ett objekt efter beskärningen tar upp en stor del av den resulterande bilden så påverkar detta den gaussiska fördelning som anpassas till bilden vid brusreduceringen. Detta kan resultera i att bilden elimineras helt eftersom hela objektet ligger innanför tröskelgränsen för brus. Detta påverkar dock endast ett fåtal bilder och de låts därför förkastas.

5.2.2 Brusreducering

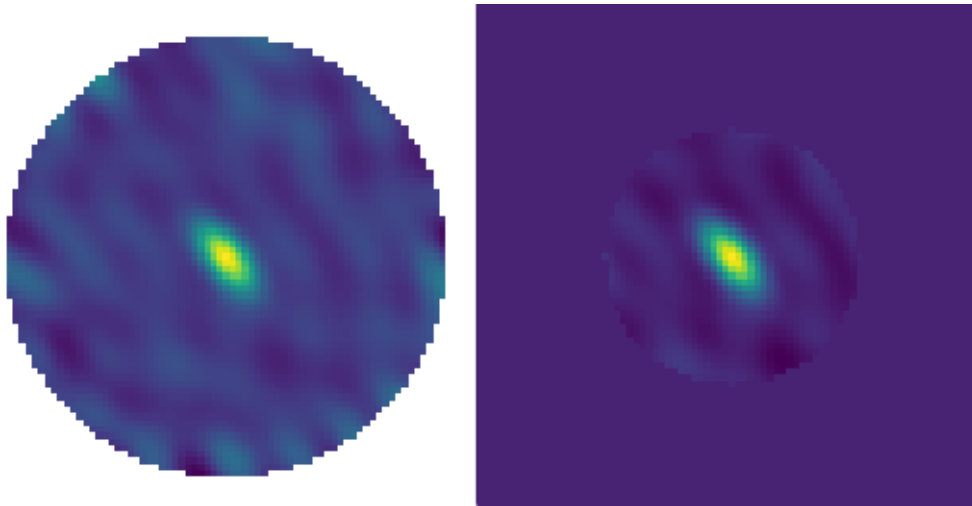
För att minimera mängden brus anpassas en gaussfunktion till filens mätvärden och sedan fokuseras bilden på de områden innehållande de högsta värdena. Denna metod valdes eftersom bruset i bilderna är relativt homogent samtidigt som objekten i bilderna ofta har tydligt högre mätvärden. För tröskelgränsen $\mu + n\sigma$ valdes $n = 5$. Alla mätvärden under det anses vara brus och inkluderas inte. Gränsen bestämdes efter flertalet tester, lägre n ger ökad risk för falska utslag på bilder där den relativa skillnaden mellan bruset och objektets mätvärden är låg och leder till att en större mängd av bakgrunden tas med. Högre n ökar risken att beskära bilderna för nära objekten och riskerar då att förstöra karakteristiska delar av objektets struktur. I figur 5.1 representeras olika val av sigma på objektet *HD163296*.



Figur 5.1: Exempel på hur olika n och tröskelvärden påverkar objektet *HD163296*.

5.2.3 Selektiv bildeliminering

Bilder där alla pixelvärden är mindre än $\mu + 7 \cdot \sigma$ avlägsnas helt innan förstoring, vilket exkluderar majoriteten av bilderna innehållande endast brus. Bilderna utan tydliga objekt som inte exkluderas av detta skiljer sig från övriga behandlade bilder och bör kategoriseras annorlunda i maskininlärningen. Metoden riskerar att ta bort observationer med starka centralkällor som upptar stor del av bilden. Ett exempel på en sådan observation illustreras i figur 5.2. I detta fall påverkas normalfördelningen kraftigt av centralkällan och både μ och σ blir högre än representativt för bruset. En möjlig framtida lösning är att anpassa en normalfördelning till flera olika bildområden. Sedan kan de delar med stora μ , som troligen innehåller objektet, exkluderas från analysen.

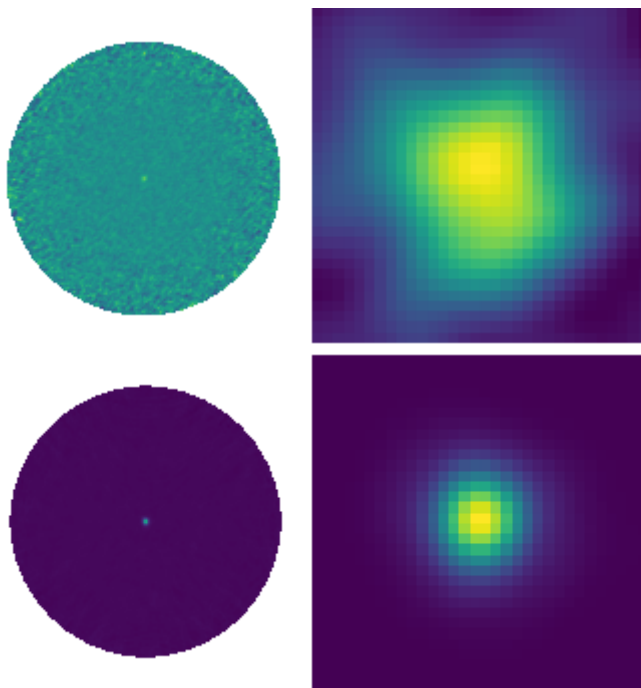


Figur 5.2: Exempel på bild ur ALMA-arkivet där den starka centralkällan upptar en större andel. Till vänster visas originalbilden, till höger den beskärda bilden.

5.2.4 Förstoring

Den slutgiltiga varianten av förstoringfunktionen väljer framgångsrikt det objekt med det maximala värdet i bilden. Majoriteten av bilderna ser ut likt figur 4.1a, 4.1b eller 4.1c där förstoringen fungerar som förväntat. De slutgiltiga objekten får liknande storlek i bilden, i synnerhet vid mindre n . Till skillnad från tidigare förstoring algoritmer förstoras ringformade objekt och skivor innehållande ringar generellt likt förväntat, se 4.1e. Objektens storlek i bilden tycks påverka klustringen, eftersom objekten i varje kluster har liknande storlek. Detta syns exempelvis på resultatbilderna 4.9 och 4.10.

I brusiga bilder som fortfarande innehåller ett objekt identifieras objektet korrekt i kombination med brusreduceringen, se figur 4.1f. Detta givet att den starkast lysande pixeln inom den beskärda cirkeln tillhör objektet. Sådana bilder beskärs närmare objektet än övriga då bruset ger en högre varians, vilket innebär större $\mu + n \cdot \sigma$. Detta påverkar troligen klustringen och kan undvikas genom utveckling av brusreduceringmetoden. En jämförelse mellan denna typ av bild och en mindre brusig bild ses i figur 5.3. Förmågan att hitta hela objekt i brusiga bilder beror kraftigt på storleken på n . För $n = 3$ fångas större andel av objektet, med ökad risk att behålla brus som påverkar klustringen. Vid $n = 7$ ökar risken att delar av objektet betraktas som brus och försvinner, vilket ses tydligt i figur 4.1f.



Figur 5.3: Jämförelse mellan bildbehandlingen på en ovanligt brusig bild och en mindre brusig bild. Till vänster visas originalbilden och till höger den slutgiltiga bilden.

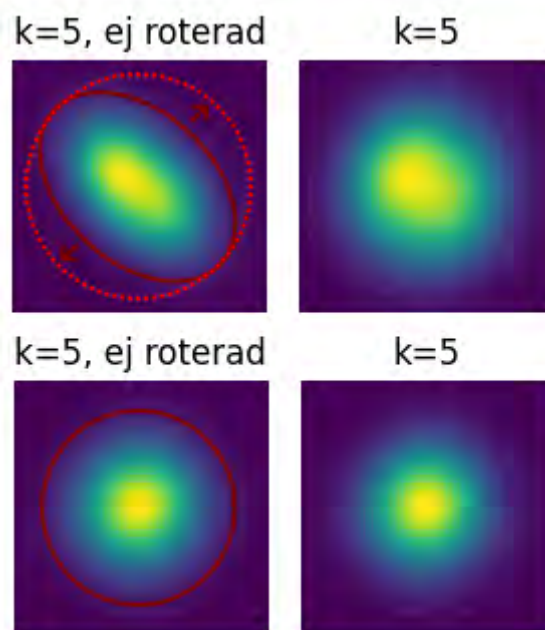
Vid högre värden på n splittras vissa ringar som är svaga i ena änden, se figur 4.1d. De svagare delarna av ringen innehåller värden nära brusnivån och riskerar därmed att bli bortskurna av förstöringsalgoritmen. För att förstora sådana ringar jämnt behöver en ny algoritm implementeras. Klustringen bör inte påverkas märkvärt eftersom denna typ av bilder för närvarande hamnar i samma kluster, vilket är önskvärt.

5.2.5 Skaländring och rotation

Hur rotationsfunktionen influerar bildernas indelning i kluster analyseras med hjälp av tabell 4.2. Här syns att bilderna med rotationsfaktor $k_R \geq \frac{3}{2}$ eller $k_R \leq \frac{2}{3}$ placerades mer utspritt i klustrena när de roterades till skillnad från när de inte roterades. Bilderna från datamängden med rotation, D , hamnade i fler kluster än de från datamängden utan rotation, C , med alla tre prövade maskininlärningsmodeller. Fördelningen tyder på att rotationsfunktionen minskar påverkan från skivornas rotation på slutresultatet, vilket är syftet med funktionen. Trots rotationen hamnar relativt många av dessa bilder i samma kluster, framförallt med modellen som använder DBSCAN och t-SNE. Detta kan möjligen bero på de problem beskrivna nedan.

Ett noterbart problem med skaländringen och rotationen är att vissa delar av diskarna blir ljusare än de hade varit vid en observation framifrån. En roterad disk får starkare ljus per pixel än en disk sedd framifrån, eftersom mer damm och gas befinner sig i varje pixel. När disken roteras kommer dessa högre värden fortsatt synas i hela disken, vilket resulterar i att disken blir något ljusare än den borde vara.

Ytterligare ett problem är att roterade objekt tenderar att bli något större i bilden än icke-roterade. Detta beror på att längden på den längre diagonalen används som slutgiltig diameter av objektet. Denna diagonal är längre än vad sidan hade varit, eftersom bilden är kvadratisk. En illustration av detta problem ses i figur 5.4. För att lösa detta problem hade pixlarna i en mindre, mer representativ, cirkel kunnat plockas från hela objektet på ett liknande sätt som rotationen genomförs i nuläget. En alternativ lösning är att lägga till en ram runtomkring bilden före skaländringen, vars tjocklek baseras på rotationsfaktorn k_R .



Figur 5.4: Illustration av hur rotationen ger större objekt i slutresultatet.

5.3 Maskininlärning

I detta avsnitt diskuteras valet av olika metoder och parametrar som har använts vid implementationen av maskininlärningsmodellen.

5.3.1 Attributextrahering

Med tanke på resultaten anses attributextrahering med DCT vara mest lämplig för vår data. Klustringsresultaten där CNN eller HOG användes anses vara märkbart sämre. Kluster med HOG ger, illustrerat i figur 4.5, inte rimliga kluster eftersom många liknande bilder placeras i olika kluster och helt olika bilder hamnar i samma.

Resultaten från DCT och pixelvärden i figur 4.3 och 4.4 är i detta fall identiska, med samma klustringsresultat och silhuettvärde. Vid tester på mindre datamängder verkade DCT ge bättre resultat. Därför användes huvudsakligen DCT för vidare tester. Även förmågan att lättare hantera större bilder gör DCT mer lämplig för

framtida arbeten. Pixelvärdesmetoden visade sig dock ge liknande resultat och är därmed också lämplig för den här datamängden.

På den större datamängden ger även CNN någorlunda rimliga kluster. Klustrena från figur 4.6 bedöms också rimliga då liknande observationer ser ut att till stor del hamna i samma kluster. Överlag anses dock denna klustring sämre än med pixel- eller DCT-attributextrahering. Ett tydligt exempel är klustering av de halvmåneformade observationerna, exempelvis de två översta bilderna i det tredje klustret i figur 4.6. Med CNN placeras halvmånebilderna ofta i kluster tillsammans med andra typer av bilder, detta sker däremot inte med klustermetoder som använt DCT eller pixelvärden. Silhuettvärdena i tabell 4.3 visar också tydligt att CNN överlag leder till sämre kluster, men betydligt bättre än för HOG.

Vid användning av silhuettvärden för analys stämmer silhuettvärdena bra överens med den visuella bedömningen av klustrena. Användning av HOG ger märkbart sämre kluster än övriga metoder vilket även silhuettvärdena indikerar. Silhuettvärdena visar också att CNN är bättre än HOG men sämre än DCT och pixelvärden vilket stämmer bra överens med den subjektiva bedömningen.

5.3.2 Dimensionsreducering

Även om både PCA och t-SNE har implementerats för dimensionsreducering beslutades endast användning av t-SNE. Beslutet grundades på att t-SNE gav lovande resultat och PCA skulle behöva justeras efter datamängden ytterligare för att ge rimliga resultat. Då modeller i andra steg av projektet krävde parameterjustering prioriterades inte testning av PCA.

Figur 4.7 och 4.8 visar hur t-SNE reducerar attributen från 1024 till 3 dimensioner. Fallet med 10 kluster visar tydligare hur t-SNE fungerar och att datan i 2 av 3 plan delas på så sätt att den efterföljande klustringsalgoritmen, i detta fall k -means, lättare kan genomföra en rimlig klustring. De valda parametrarna verkar därmed fungera för datamängden. Figuren med 70 kluster visar t-SNEs indelning under samma förutsättning som för de övriga resultaten. Figuren är svårare att tolka eftersom det finns 70 olika kluster och därmed 70 olika färger i figuren.

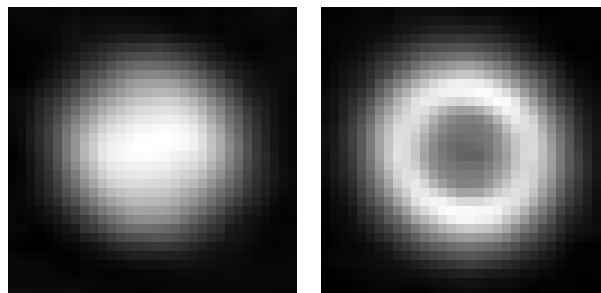
För att mer konkret se hur t-SNE påverkar klustringen kan figur 4.9 jämföras med 4.10 och figur 4.11 med 4.12. Särskilt för DBSCAN anses att kluster framtagna med t-SNE är betydligt bättre. Utan t-SNE placeras majoriteten av bilderna i kluster 0 och i brusklustret -1 . Många av bilderna i brusklustret är dessutom mycket lika bilder i de övriga klustrena och bör rimligtvis placeras där istället. Med t-SNE läggs fortfarande många bilder i brusklustret, men modellen ger överlag många fler kluster bestående av bilder på olika typer av skivor.

Vid betraktning av tabell B.4 för DBSCAN med t-SNE observeras exempelvis att bilder på objekt HD 142527, som i arkivet finns sparad under objektnamn HD_142527, Hd_142527 och hd142527, oftast placeras i kluster 16 med vissa observationer i klus-

ter 61, 26 och 97. För motsvarande klustring i tabell B.3 för DBSCAN utan t-SNE läggs dessa observationer till stor del i brusklustret -1 , men även utspritt i kluster 4, 6, 12, 13 och 26. Denna spridning för observationer av samma objekt tyder på att klustringen blir bättre då t-SNE används.

Det finns skillnader även för klustring genomförd med k -Means. En direkt jämförelse mellan figurerna 4.9 och 4.10 visar att strukturen på klustrena varierar mycket. Utan t-SNE varierar klustrenas längd kraftigt. Flertalet kluster har endast en observation och andra har över 20 stycken. Med t-SNE erhålls istället kluster av enhetligare storlek. Genom att analysera olika objekts placering i tabellerna B.1 och B.2 observeras att olika observationer av samma objekt ofta placeras i samma kluster både med och utan t-SNE.

Viktigt att betona är att det i vissa fall kan vara missvisande att enbart studera om samma objekt placeras i samma kluster. Ett exempel på detta är planetsystemet V4046 Sagittarii, sparad som V4046_Sgr, som har observerats i 6 separata bilder i datamängden. Både med och utan t-SNE placeras bilderna i separata kluster men genom att analysera dem kan detta förklaras. Figur 5.5 visar 2 av dessa 6 observationer och hur olika de är. Att dessa två bilder hamnar i olika kluster är förväntat eftersom det finns andra observationer av objekt som liknar de två bilderna mer. Att bilderna är väldigt olika trots att de föreställer samma objekt beror på att objekten har observerats vid olika våglängder och med olika konfigurationer av ALMA. Exempelvis kan emissionsnivåerna i olika spektrum skilja sig åt beroende på storleken av det stoft som skivan till största del består av. Att observationerna kan bli mycket olika är därmed inte förvånande.



Figur 5.5: Exempel på 2 olika observationer av V4046 Sagittarii. Bilderna är tagna från datamängd D.

5.3.3 Klustringsmodeller

I denna del av diskussionen analyseras valet av maskininlärningsmetoder, deras effektivitet och lämplighet för uppgiften att klustra bilder på protoplanetära skivor. Initialt övervägdes samtliga implementerade klustringstekniker, som beskrivits utförligt i avsnittet om klusteranalys 2.3.3. För att jämföra de olika modellerna baseras bedömningen främst på resultatbilderna som återfinns i avsnitt 4.3.3 och 4.3.4. Testerna visar, enligt utvärderingsmetoden från avsnitt 3.4, att figur 4.10 med $k = 70$

resulterar i rimligast kluster för den valda datamängden.

Klustringen med k -means anses bli bättre än den med DBSCAN eftersom storleken på klustrena är likartade. I både figur 4.11 och 4.12 noteras att brusklustret blir mycket långt och utan t-SNE blir även kluster 0 långt. K -means leder istället till kluster av jämnare storlek vilket anses vara önskvärt. Även med t-SNE finns det många bilder i brusklustret som rent visuellt är mycket lika bilder i flera andra kluster och som rimligtvis skulle ha kunnat placeras där istället.

Båda modellerna verkar hantera placeringen av olika observationer på samma objekt på liknande sätt. Bilderna placeras oftast i samma kluster för båda modellerna och många av situationerna där det inte sker kan förklaras med den iakttagelsen som beskrevs under föregående avsnitt vid figur 5.5. Överlag ger DBSCAN klustringsresultat av en annan struktur som inte nödvändigtvis är sämre. Dock anses de kluster som DBSCAN genererar mindre lämpliga för projektets syfte.

Kravet att k -means behöver ett bestämt k -värde belyser ett stort problem med algoritmen, dess oförmåga att autonomt bestämma ett optimalt antal kluster. Problemet är kritiskt givet variationen i datamängderna som behandlas och för att programmet ska kunna hantera andra datamängder i framtiden utan att parametrar ska behöva justeras. Denna insikt ledde till att alternativet DBSCAN utforskades mer, som erbjöd automatiskt val av antal kluster och kluster med godtagbar kvalitet. Klustrena som DBSCAN genererar är inte lika rimliga som de med k -means men indikerar det antal kluster som är rimligast för datamängden. En möjlig användning av DBSCAN är därmed att få ett ungefärligt värde på parametern k genom att ansätta det till antalet kluster som DBSCAN genererar.

I DBSCAN bestämdes parametrarna ε och "*min_samples*" genom tester där kluster genererades med olika kombinationer av parametrar. I testerna valdes kombinationer av parametrar i enlighet med utvärderingsmetoden i 3.4, där de mest gynsamma parametrarna ledde till att ett objekt mestadels placerades i samma kluster. Ett exempel på sådana kluster är de i figur 4.12 med $\varepsilon = 1$ och "*min_samples*" = 2.

GMM, som beskrivs i avsnittet klustring 2.3.3.2, övervägdes också för att dra nytta av dess förmåga att bestämma klusterantalet automatiskt med BIC. Den automatiska bestämningen av antal kluster för GMM resulterar i avsevärt färre kluster än för de övriga modellerna, vilket kraftigt påverkar kvalitén på klustrena. Ett exempel på klustringsresultat från GMM återfinns i Appendix A, figur A.1.

5.4 Jämförelse med liknande projekt

Ett liknande projekt som genomförts tidigare inom maskinlärning och astronomi är Fagrell, Kollberg, Rasmussen m. fl. [6]. En stor likhet mellan studierna är att bilder på protoplanetära skivor från ALMA-arkivet analyseras. För att analysera och klassificera de astronomiska bilderna använder även Fagrell, Kollberg, Rasmussen m. fl. [6] maskinlärning, särskilt CNN. Även syftet med deras studie kan kopplas till vårt

projekt. Detta eftersom syftet med projekten är att identifiera och gruppera protoplanetära skivor baserat på observationer för att få en ökad förståelse inom ämnet.

Ett specifikt syfte som skiljer arbetet av Fagrell, Kollberg, Rasmussen m. fl. [6] från vårt är en tillämpning av maskininlärning för att hitta potentiella indikationer på stoft i protoplanetära skivvindar. Istället fokuserar vårt projekt på en bred analys av ALMA-data för att försöka identifiera mönster eller egenskaper hos protoplanetära skivor. Medan Fagrell, Kollberg, Rasmussen m. fl. [6] använder sig utav övervakad inlärning genom att träna faltningsnätverk, bygger vårt projekts maskininlärningsmodell på oövervakad inlärning. Implementationen av modellerna och utvärderingsmetoderna skiljer sig för att anpassa maskininlärningen för respektive projekt. Båda projekten syftar till att förenkla navigering i ALMA-arkivet, men de specifika resultaten och slutsatserna skiljer sig åt på grund av olika tillvägagångssätt.

6

Slutsats

I detta projekt har en maskininlärningsmodell implementerats för att gruppera observationer från ALMA-arkivet med syftet att automatisera och förenkla navigeringen i arkivet. Genom tester med olika maskininlärningsmetoder och bildbehandlingsalgoritmer har betydande resultat uppnåtts.

Det har identifierats att bildprocessering är kritisk för att korrekt klustra astronomiska bilder och projektet har utforskat olika tekniker för reducera påverkan från exempelvis brus, objektets position i bilden och dess inklanation i himmelsplanet. Lokalisering med tröskelnivå $\mu + 5\sigma$ samt behandling av objektens inklanation i himmelsplanet visade sig vara mest lämpligt för att förbereda bilderna inför efterföljande maskininläring.

Testerna med olika maskininlärningsalgoritmer inom både attributextrahering, dimensionsreducering och klustring har lett till intressanta resultat gällande klusteranalys av astronomiska bilder. Både för- och nackdelar med olika metoder och parametrar har observerats, vilket har bidragit till en djupare förståelse för hur algoritmerna fungerar på den valda datamängden. Specifikt har kombinationen av attributextrahering med DCT, dimensionsreducering med t-SNE och klustring med k -means där $k = 70$, visat sig vara den mest framstående med datamängden.

Det finns betydande potential för att förbättra och vidareutveckla den nuvarande modellen. Framförallt bör fokus ligga på att förbättra brusreduceringen, utveckla mer avancerade algoritmer för bildförstoring och hantering av inklinations samt justera parametrar på t-SNE och klusteralgoritmerna ytterligare. För alla steg i maskininlärningsmodellen finns det även fler algoritmer som kan implementeras och utvärderas. Dessutom bör framtida studier inkludera mer omfattande tester med varierande bildkvaliteter för att säkerhetsställa att modellen är robust över ett bredare spektrum av data.

Sammanfattningsvis förmodas detta projekt bidra till att automatisera och förbättra grupperingen av astronomiska bilder, vilket förväntas gynna forskare och användare genom att öka tillgängligheten och användbarheten av data i ALMA-arkivet. Fortsatt arbete inom detta område har stor potential att främja fortsatt forskning inom astronomi och bildanalys.

Litteratur

- [1] T. Birnstiel, C. P. Dullemond, Z. Zhu m. fl., "The Disk Substructures at High Angular Resolution Project (DSHARP). V. Interpreting ALMA Maps of Protoplanetary Disks in Terms of a Dust Model [fotografi]," *apjl*, årg. 869, nr 2, L45, s. L45, dec. 2018. DOI: [10.3847/2041-8213/aaf743](https://doi.org/10.3847/2041-8213/aaf743). arXiv: [1812.04043](https://arxiv.org/abs/1812.04043) [[astro-ph.SR](https://arxiv.org/abs/1812.04043)].
- [2] ALMA-observatoriet, *How ALMA Works | ALMA*. URL: <https://www.almaobservatory.org/en/about-alma/how-alma-works/>.
- [3] ALMA, *ALMA Science Archive*. URL: <https://almascience.nrao.edu/aq/>.
- [4] Ronald D Ekers, "Big and Small," i *Proceedings of Accelerating the Rate of Astronomical Discovery*, Rio de Janeiro, 2010.
- [5] U. Lebreuilly, P. Hennebelle, T. Colman m. fl., "Protoplanetary Disk Birth in Massive Star-forming Clumps: The Essential Role of the Magnetic Field," *The Astrophysical Journal Letters*, årg. 917, nr 1, s. L10, aug. 2021, ISSN: 2041-8205. DOI: [10.3847/2041-8213/ac158c](https://doi.org/10.3847/2041-8213/ac158c).
- [6] P. Fagrell, J. Kollberg, E. Rasmussen, E. R. Axelsson, O. Svensson och A. Ybring, "Finding Potential Detections of Dust in Protoplanetary Disk Winds Using Machine Learning to Filter, Classify and Present ALMA-data Bachelor's thesis in Space, Earth and Environment," tekn. rapport, 2023. URL: www.chalmers.se.
- [7] A. Fox och R. Davé, "Gas Accretion onto Galaxies," tekn. rapport. URL: <http://www.springer.com/series/5664>.
- [8] ALMA-observatoriet, *About ALMA, at first glance | ALMA*. URL: <https://www.almaobservatory.org/en/about-alma/>.
- [9] Europeiska Sydobservatoriet, *ALMA - Atacama Large Millimeter/submillimeter Array | ESO*. URL: <https://www.eso.org/public/teles-instr/alma/>.
- [10] Europeiska Sydobservatoriet, *Interferometry | ESO*. URL: <https://www.eso.org/public/teles-instr/technology/interferometry/>.
- [11] J. A. Högbom, "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines," *aaps*, årg. 15, s. 417, juni 1974, Provided by the SAO/NASA Astrophysics Data System. URL: <https://ui.adsabs.harvard.edu/abs/1974A&AS...15..417H>.
- [12] ALMA Basics, *ALMA Basics — ALMA Science Portal at ESO*. URL: <https://almascience.eso.org/about-alma/alma-basics>.
- [13] A. Ahmadi och A. Hacar. "ALminer: ALMA archive mining and visualization toolkit." (2021), URL: <https://ascl.net/code/v/2971>.
- [14] T. Jaffe, *FITS Support Office*, 2023. URL: https://fits.gsfc.nasa.gov/fits_home.html.

- [15] T. Jaffe, *FITS Introduction*, 2014. URL: https://fits.gsfc.nasa.gov/fits_overview.html.
- [16] F. Guglielmetti, D. Petry, B. Mason m. fl., "ALMA QA2 Data Products for Cycle 10, Version 1.1," ALMA, Report, 2023.
- [17] P. C. Cortes, C. Vlahakis, A. Hales m. fl., "ALMA Technical Handbook," ALMA, Technical Report ALMADoc. 10.3, version 1.0, 2023. DOI: [10.5281/zenodo.4511521](https://doi.org/10.5281/zenodo.4511521).
- [18] IBM, *What is Unsupervised Learning? | IBM*. URL: <https://www.ibm.com/topics/unsupervised-learning>.
- [19] Ryszard S. Choras, "Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems," *INTERNATIONAL JOURNAL OF BIOLOGY AND BIOMEDICAL ENGINEERING*, årg. 1, nr 1, 2010. URL: <https://www.naun.org/main/NAUN/bio/bio-2.pdf>.
- [20] C. Zheng och D. W. Sun, "Image Segmentation Techniques," *Computer Vision Technology for Food Quality Evaluation*, s. 37–56, 2008. DOI: [10.1016/B978-012373642-0.50005-3](https://doi.org/10.1016/B978-012373642-0.50005-3).
- [21] M. Lorentzon, *Feature extraction for image selection using machine learning*, Linköping, 2017. URL: <https://www.diva-portal.org/smash/get/diva2:1151145/FULLTEXT01.pdf>.
- [22] *Russian Blue Cat Napping [fotografi]*, Britannica ImageQuest, Hämtad Mars 1, 2024, 2024. URL: https://quest-eb-com.eu1.proxy.openathens.net/images/139_3822257.
- [23] IBM, *What are Convolutional Neural Networks? | IBM*. URL: <https://www.ibm.com/topics/convolutional-neural-networks>.
- [24] IBM, *What is a Neural Network? | IBM*. URL: <https://www.ibm.com/topics/neural-networks>.
- [25] I. Nunes da Silva, D. Hernane Spatti, A. Andrade Flauzino, L. Helena Bartocci Liboni och S. F. dos Reis Alves, *Artificial Neural Networks A Practical Course*. 2016, vol. 1, ISBN: 978-3-319-43162-8. URL: <https://doi-org.proxy.lib.chalmers.se/10.1007/978-3-319-43162-8>.
- [26] K. O'Shea och R. Nash, *An Introduction to Convolutional Neural Networks*, 2015. arXiv: [1511.08458 \[cs.NE\]](https://arxiv.org/abs/1511.08458).
- [27] L. van der Maaten, E. Postma och H. Herik, "Dimensionality Reduction: A Comparative Review," *Journal of Machine Learning Research - JMLR*, årg. 10, jan. 2007. URL: https://www.researchgate.net/publication/228657549_Dimensionality_Reduction_A_Comparative_Review.
- [28] L. Van Der Maaten och G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, årg. 9, s. 2579–2605, 2008.
- [29] H. P. Kriegel, E. Schubert och A. Zimek, "The (black) art of runtime evaluation: Are we comparing algorithms or implementations?" *Knowledge and Information Systems*, årg. 52, nr 2, s. 341–378, aug. 2017, ISSN: 02193116. DOI: [10.1007/S10115-016-1004-2/FIGURES/7](https://doi.org/10.1007/S10115-016-1004-2/FIGURES/7). URL: <https://link.springer.com/article/10.1007/s10115-016-1004-2>.
- [30] M. Z. Rodriguez, C. H. Comin, D. Casanova m. fl., "Clustering algorithms: A comparative approach," *PLoS ONE*, årg. 14, nr 1, jan. 2019, ISSN: 19326203. DOI: [10.1371/JOURNAL.PONE.0210236](https://doi.org/10.1371/JOURNAL.PONE.0210236). URL: <https://www.researchgate.net>.

- [net/publication/311925975_Clustering_Algorithms_A_Comparative_Approach](https://arxiv.org/abs/1908.08724).
- [31] D. A. Reynolds, "Gaussian mixture models," *Encyclopedia of biometrics.*, årg. 741, s. 659–663, 2009.
- [32] M. Ester, H.-P. Kriegel, J. Sander och X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," 1996. URL: www.aaai.org.
- [33] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, årg. 20, nr C, s. 53–65, nov. 1987, ISSN: 0377-0427. DOI: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- [34] O. Arbelaitz, I. Gurrutxaga, J. Muguerza, J. M. Pérez och I. Perona, "An extensive comparative study of cluster validity indices," *Pattern Recognition*, årg. 46, nr 1, s. 243–256, jan. 2013, ISSN: 0031-3203. DOI: [10.1016/J.PATCOG.2012.07.021](https://doi.org/10.1016/J.PATCOG.2012.07.021).
- [35] ttnphns (<https://stats.stackexchange.com/users/3277/ttnphns>), *Comparing silhouette scores between different datasets (having different number of variables)*, Cross Validated, URL:<https://stats.stackexchange.com/q/223341> (version: 2016-07-12). eprint: <https://stats.stackexchange.com/q/223341>. URL: <https://stats.stackexchange.com/q/223341>.
- [36] F. Pedregosa, G. Varoquaux, A. Gramfort m. fl., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, årg. 12, s. 2825–2830, 2011.
- [37] Keras 3 API documentation, *VGG16 and VGG19*, 2023. URL: <https://keras.io/api/applications/vgg/>.
- [38] P. Virtanen, R. Gommers, T. E. Oliphant m. fl., "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, årg. 17, s. 261–272, 2020. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).
- [39] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias m. fl., "scikit-image: image processing in Python," *PeerJ*, årg. 2, e453, juni 2014, ISSN: 2167-8359. DOI: [10.7717/peerj.453](https://doi.org/10.7717/peerj.453). URL: <https://doi.org/10.7717/peerj.453>.
- [40] G. Schwarz, "Estimating the Dimension of a Model," *Annals of Statistics*, årg. 6, nr 2, s. 461–464, mars 1978, ISSN: 0090-5364. DOI: [10.1214/AOS/1176344136](https://doi.org/10.1214/AOS/1176344136). URL: <https://projecteuclid.org/journals/annals-of-statistics/volume-6/issue-2/Estimating-the-Dimension-of-a-Model/10.1214/aos/1176344136.full><https://projecteuclid.org/journals/annals-of-statistics/volume-6/issue-2/Estimating-the-Dimension-of-a-Model/10.1214/aos/1176344136.short>.

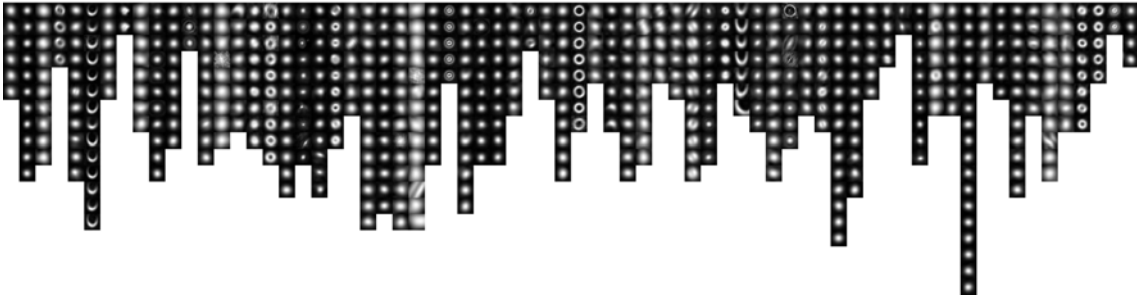
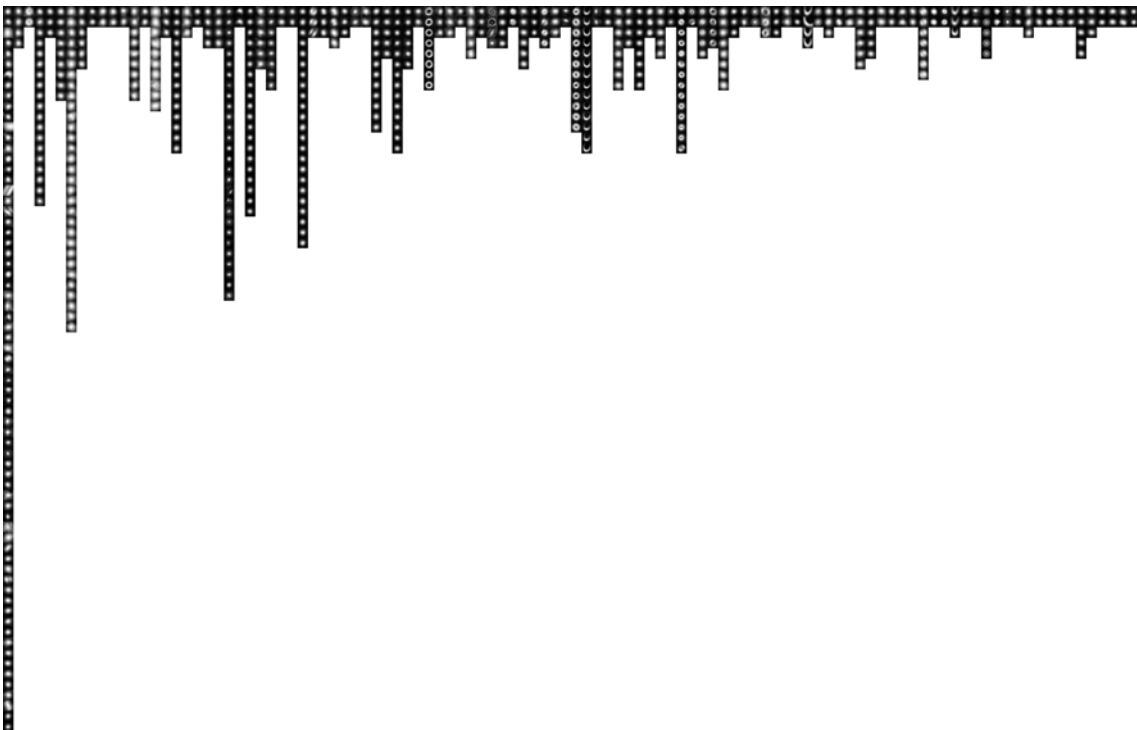
A

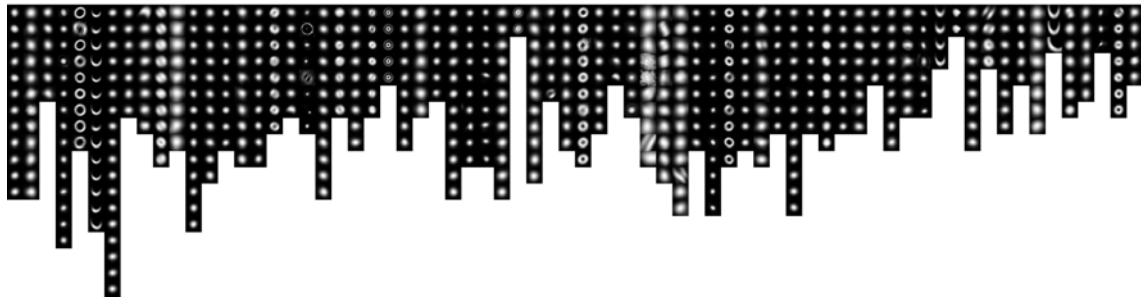
Övriga resultatbilder

I detta avsnitt presenteras ett brett urval av övriga klustringsresultat som tagits fram med olika parametrar och val av modeller. Dessa resultat är överlag sämre än de som presenteras i resultatavsnittet men belyser hur valet av olika parametrar och modeller påverkar klustringarna. För alla klustringar i detta avsnitt har t-SNE använts, men parametrar för bildbehandling, k -Means DBSCAN varierar. Resultatbilderna i sin ursprungliga form och storlek finns även tillgängliga i projektets GitHub. Se appendix [C](#) för mer information om detta.



Figur A.1: Exempel på klustering med GMM och t-SNE. Notera att antalet kluster är betydligt lägre än för de övriga modellerna med bara 6 kluster. Bilderna initialt $\sigma = 5$ roterade, sedan $\sigma = 7$ utan oskarpa.

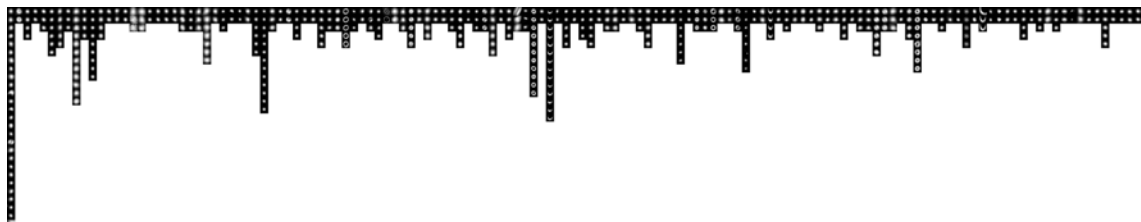
(a) k -Means, $k = 70$.(b) k -Means, $k = 120$.(c) DBSCAN, $\epsilon = 0.7$, 'min_samples'=2. Brusklustret är egentligen mycket längre och har till stor del klippts bort likt figur 4.11 för att få plats i rapporten.(d) DBSCAN, $\epsilon = 1$, 'min_samples'=2.**Figur A.2:** Klustringsresultat med olika klustringsmodeller och parametrar. Bilderna initialt $\sigma = 3$ roterade, sedan $\sigma = 7$ utan oskarpa.



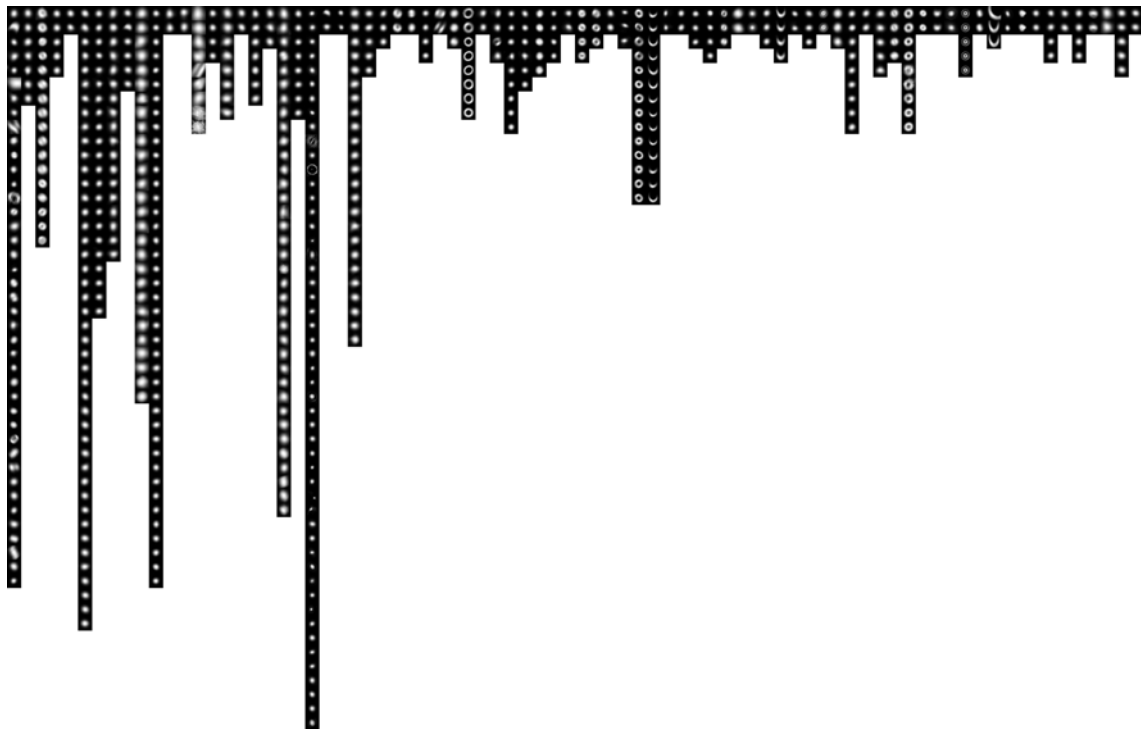
(a) k -Means, $k = 70$.



(b) k -Means, $k = 120$.

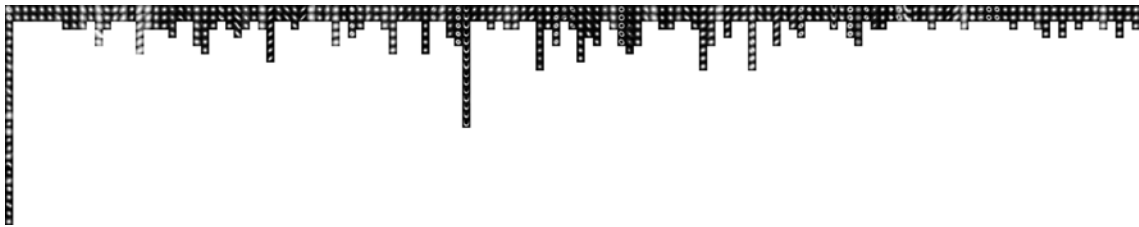
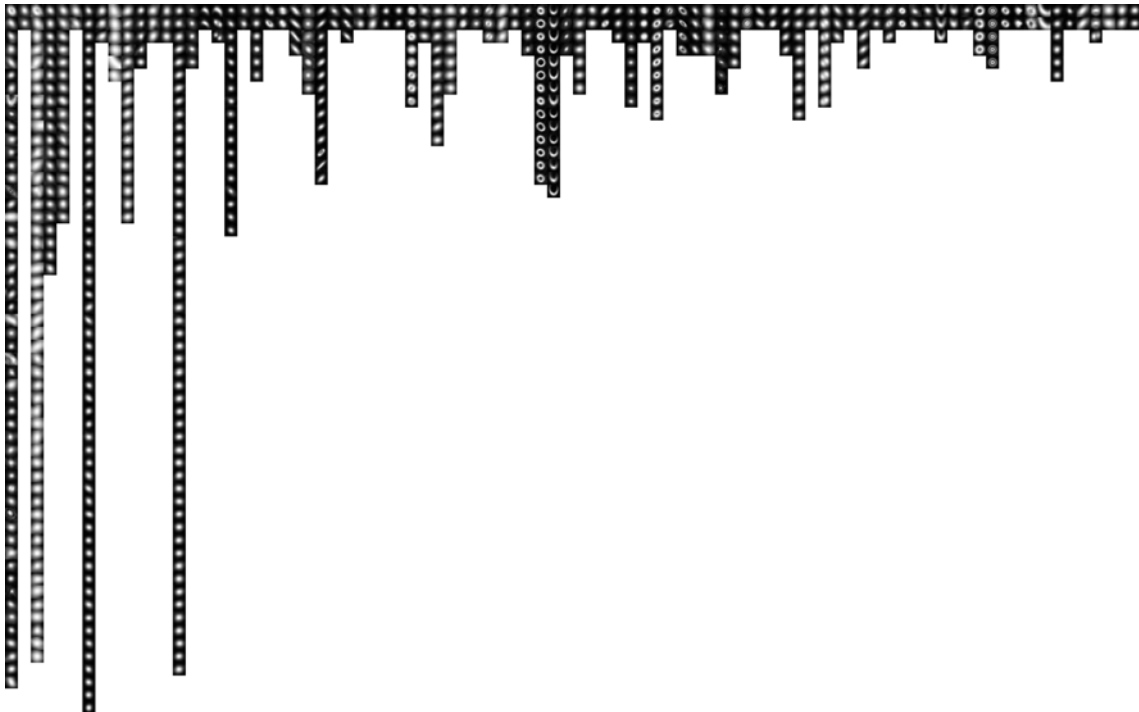


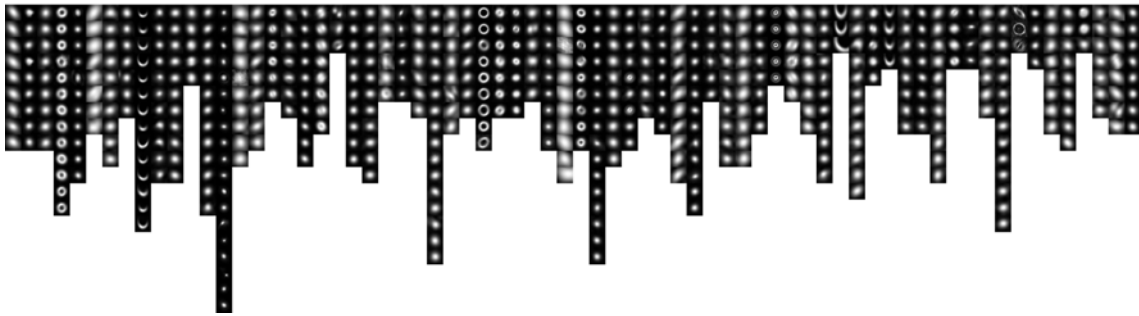
(c) DBSCAN, $\epsilon = 0.7$, 'min_samples'=2. Brusklustret är egentligen mycket längre och har till stor del klippts bort likt figur 4.11 för att få plats i rapporten.



(d) DBSCAN, $\epsilon = 1$, 'min_samples'=2.

Figur A.3: Klustringsresultat med olika klustringsmodeller och parametrar. Bilderna initialt $\sigma = 3$ roterade, sedan $\sigma = 7$ med oskarpa.

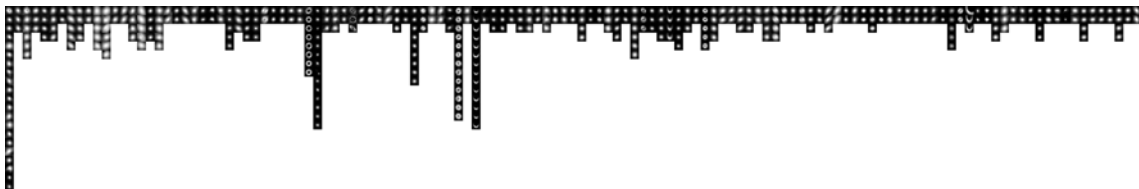
(a) k -Means, $k = 70$.(b) k -Means, $k = 120$.(c) DBSCAN, $\epsilon = 0.7$, 'min_samples'=2. Brusklustret är egentligen mycket längre och har till stor del klippts bort likt figur 4.11 för att få plats i rapporten.(d) DBSCAN, $\epsilon = 1$, 'min_samples'=2.**Figur A.4:** Klustringsresultat med olika klustringsmodeller och parametrar. Bilderna initialt $\sigma = 5$ ej roterade, sedan $\sigma = 7$ utan oskarpa.



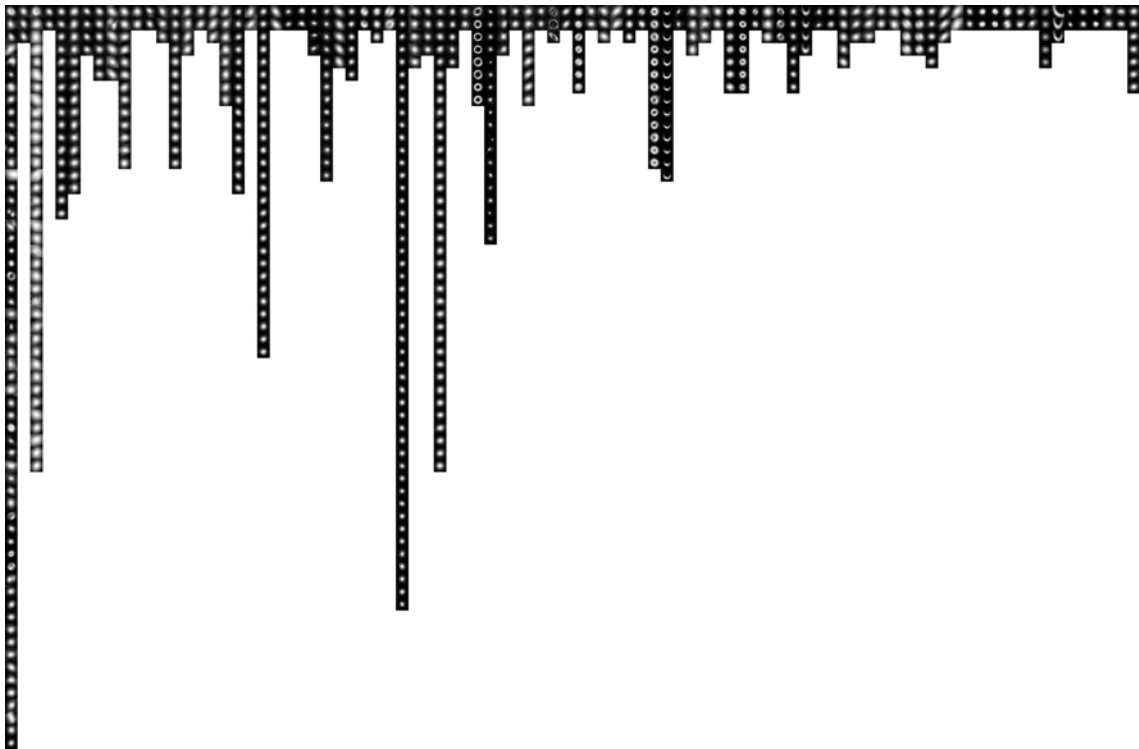
(a) k -Means, $k = 70$.



(b) k -Means, $k = 120$.

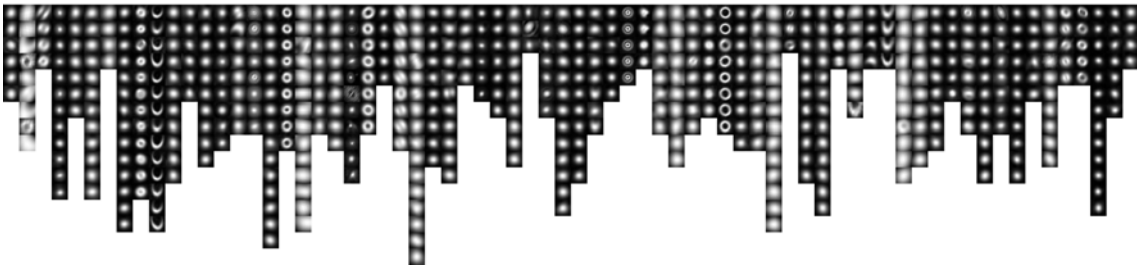
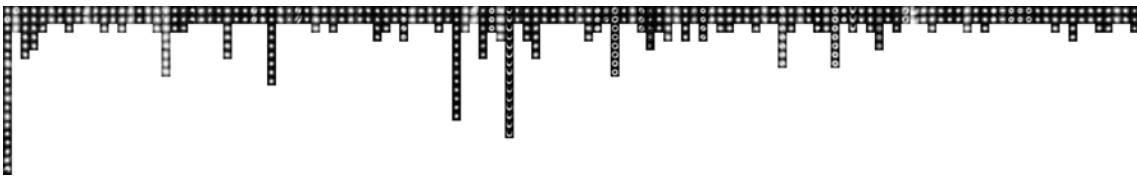
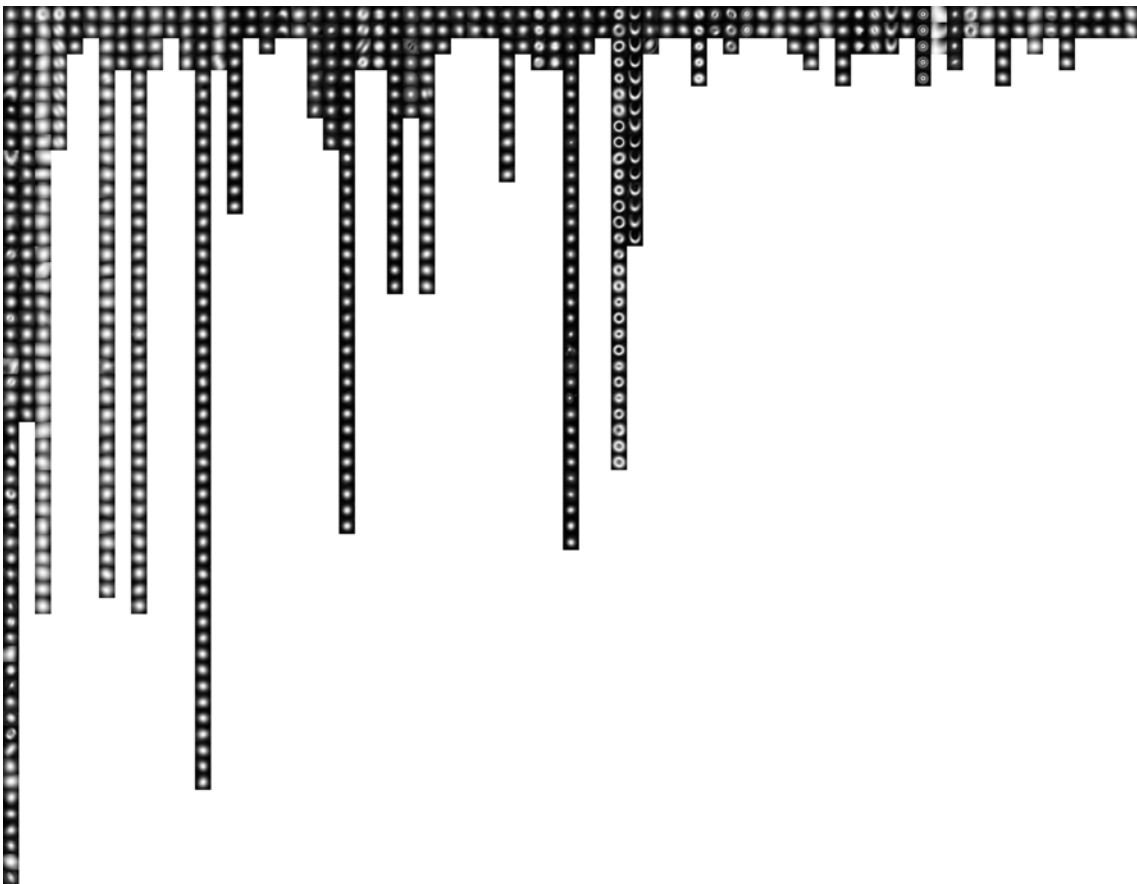


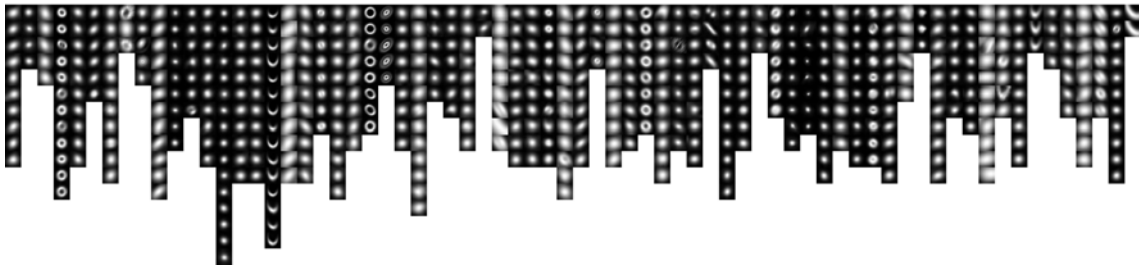
(c) DBSCAN, $\epsilon = 0.7$, 'min_samples'=2. Brusklustret är egentligen mycket längre och har till stor del klippts bort likt figur 4.11 för att få plats i rapporten.



(d) DBSCAN, $\epsilon = 1$, 'min_samples'=2.

Figur A.5: Klustringsresultat med olika klustringsmodeller och parametrar. Bilderna initialt $\sigma = 5$ roterade, sedan $\sigma = 7$ med oskarpa.

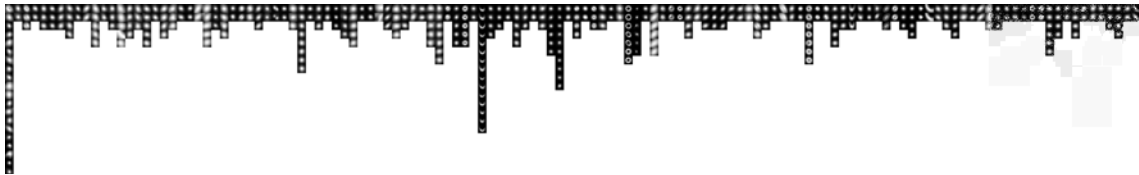
(a) k -Means, $k = 70$.(b) k -Means, $k = 120$.(c) DBSCAN, $\epsilon = 0.7$, 'min_samples'=2. Brusklustret är egentligen mycket längre och har till stor del klippts bort likt figur 4.11 för att få plats i rapporten.(d) DBSCAN, $\epsilon = 1$, 'min_samples'=2.**Figur A.6:** Klustringsresultat med olika klustringsmodeller och parametrar. Bilderna $\sigma = 7$ roterade, utan oskarpa.



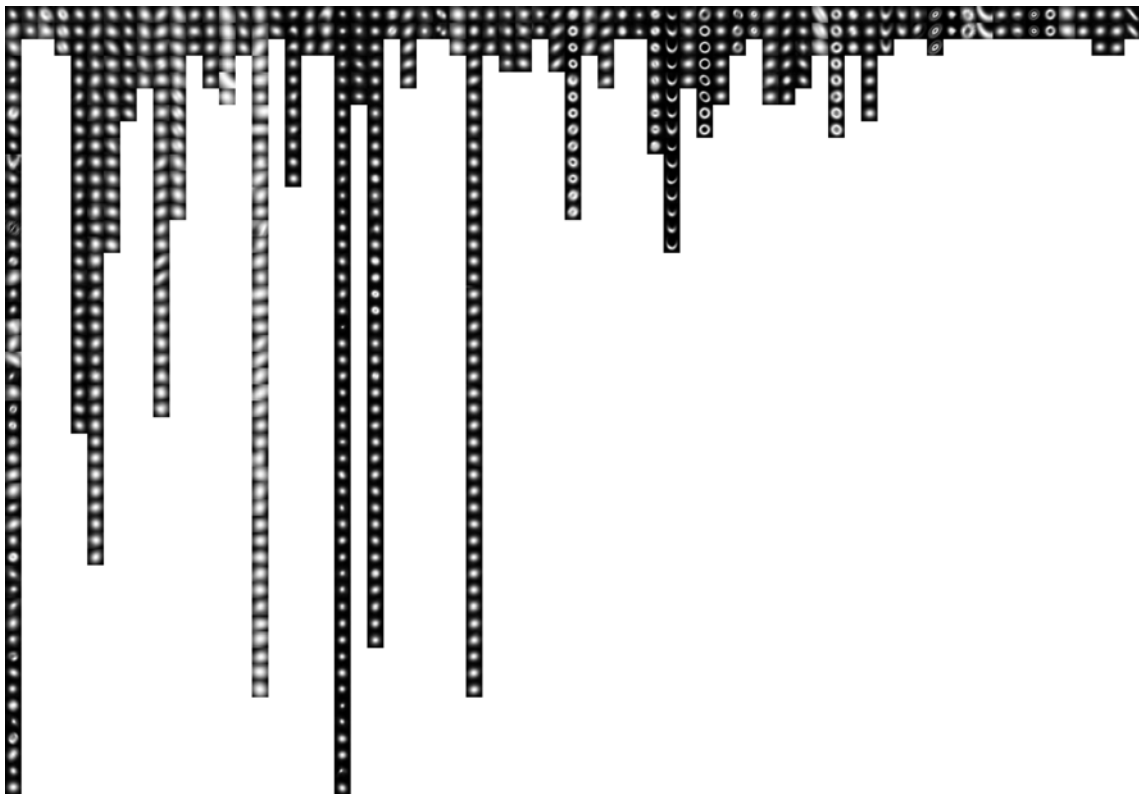
(a) k -Means, $k = 70$.



(b) k -Means, $k = 120$.



(c) DBSCAN, $\epsilon = 0.7$, 'min_samples'=2. Brusklustret är egentligen mycket längre och har till stor del klippts bort likt figur 4.11 för att få plats i rapporten.



(d) DBSCAN, $\epsilon = 1$, 'min_samples'=2. Ett kluster har delvis klippts bort för att bilden ska få plats.

Figur A.7: Klustringsresultat med olika klustringsmodeller och parametrar. Bilderna $\sigma = 7$ roterade, med oskarpa.

B

Klusterinformation

I denna bilaga presenteras information om objekt som förekommer flera gånger i vår datamängd samt deras tillhörighet till kluster från avsnitt 4.3.3 och 4.3.4. Tabell B.1 motsvarar resultaten från figur 4.9, medan tabell B.2 relaterar till figur 4.10. Likaså korresponderar tabell B.3 med figur 4.11 och tabell B.4 med figur 4.12.

Tabell B.1: *K*-means utan t-SNE

Objekt	Kluster
IM_Lup	0, 0, 9, 39, 39, 47, 69
HD_97048	0, 0, 18
TWA_7	0, 20
mwc480	0, 8, 18, 22, 22, 42, 49, 49, 49
G35.03+0.35	0, 41
serp_smm9d	0, 38
S255_IR	0, 23
TW_Hya	0, 4, 8, 9, 20, 39, 60, 63, 69
LupusIII_21	2, 42
Oph_IRS_48	2, 2, 18, 58, 62
Objekt utan namn	3, 35, 37, 43, 49, 51
Hd_142527	3, 8, 8, 21, 21
HD_142527	3, 10, 10, 10, 19, 19, 19, 19, 19, 21, 28, 51, 51, 51
hd142527	3, 3
PDS_70	5, 5
HD163296	6, 6, 56, 69
HD_163296	6, 39, 39, 59, 59
2MASS_J16042165-2130284	7, 68, 68
UX_Tau_A	7, 15
UX_Tau	7, 49
RXJ1852.3-3700	7, 7, 30, 30, 37
serp_smm4	8, 20
J16082324-1930009	8, 34
J16090075-1908526	8, 12
J15583692-2257153	8, 47
DM_Tau	8, 20, 22, 33, 35, 46, 58, 58
V4046_Sgr	9, 9, 36, 36, 46, 46
AA_Tau	9, 24, 26, 42
GM_Aur	9, 18, 18
J04362151+2351165	13, 13

B. Klusterinformation

Objekt	Kluster
LupusIII_91	13, 47
AS_209	14, 14, 14, 20
AS_205	14, 20, 24
CY_Tau	14, 69, 69
CX_Tau	14, 67
RA16_27_13.807	14, 22
HD_100453	15, 15, 42
LupusIII_1015	15, 55
pds99	15, 15
TCha	17, 57
lkca15	18, 18, 45, 45, 45, 45, 45, 45, 45
DM_tau	18, 18
RA16_31_31.243	18, 44
J04295950+2433078	18, 29
RXJ1842.9-3532	18, 25, 25, 25, 25
Serpens-FIRS1	20, 24
IRS_63	20, 42, 56
DO_Tau	20, 20
UZ_Tau	20, 20
J16113134-1838259	24, 49
2MASS_J16085553-3902339	26, 34
RA16_32_21.038	26, 58
hd169142	27, 27
HD_169142	27, 27, 29
LupusIII_19	29, 29
HD_34282	29, 30, 46, 57, 57
J16042165-2130284	29, 49
2MASS_J04220069+2657324	29, 48
RW_Aur	35, 35, 58
Wa_Oph_6	35, 35
G35.20-0.74N	35, 47
LkCa_15	36, 36, 36, 36
WaOph_6	39, 39
serp_smm9	39, 40
serp_smm2	41, 49
VSSG_1	42, 42
HN_Tau	42, 42
V710_Tau	42, 42, 54
sz84	42, 49
serp_smm11	42, 58
RY_Tau	42, 42
IRAS_04187+1927	44, 44
CIDA_1	46, 46, 49
RA16_21_31.920	47, 59
HD102647	49, 49
GO_Tau	56, 56
GSS_31	58, 58
Parsamian_21	58, 58

Objekt	Kluster
RNO_90	58, 58, 69
LupusIII_53	58, 58

Tabell B.2: *K*-means med t-SNE

Objekt	Kluster
mwc480	0, 0, 6, 10, 10, 12, 17, 28, 58
GSS_31	1, 19
AS_205	1, 51, 59
RW_Aur	1, 1, 56
Parsamian_21	1, 19
HD_169142	2, 27, 27
HD_34282	2, 3, 3, 3, 15
J04295950+2433078	2, 60
TCha	3, 16
AA_Tau	4, 32, 50, 50
serp_smm9	4, 19
Objekt utan namn	5, 5, 15, 24, 56, 66
HD_142527	5, 5, 5, 5, 5, 5, 5, 5, 5, 45, 45, 45, 45, 45
Hd_142527	5, 28, 28, 45, 45
hd142527	5, 5
DM_Tau	6, 6
Serpens-FIRS1	7, 7
serp_smm2	8, 12
LupusIII_91	8, 47
AS_209	9, 9, 9, 29
RNO_90	9, 19, 43
DM_Tau	9, 12, 16, 19, 30, 44, 44, 56
DO_Tau	9, 59
UZ_Tau	10, 59
RY_Tau	10, 52
TW_Hya	11, 13, 22, 25, 46, 48, 48, 48, 53
HD163296	11, 18, 43, 43
HD102647	12, 12
2MASS_J16085553-3902339	13, 38
J16082324-1930009	13, 28
Oph_IRS_48	14, 19, 37, 37, 41
CIDA_1	15, 44, 44
V4046_Sgr	15, 15, 16, 16, 63, 63
RXJ1842.9-3532	15, 20, 20, 20, 20
LkCa_15	16, 16, 16, 16
GO_Tau	18, 18
Wa_Oph_6	18, 18
S255_IR	18, 33
IRS_63	18, 23, 59
serp_smm11	19, 51

Objekt	Kluster
LupusIII_53	19, 30
RA16_31_31.243	21, 22
HD_97048	22, 68, 68
serp_smm4	23, 28
LupusIII_21	23, 37
J16042165-2130284	24, 34
J15583692-2257153	24, 28
J16090075-1908526	25, 34
hd169142	27, 27
RA16_27_13.807	29, 62
RA16_32_21.038	30, 38
TWA_7	33, 59
serp_smm9d	33, 55
J16113134-1838259	34, 34
2MASS_J16042165-2130284	35, 35, 67
RXJ1852.3-3700	35, 35, 66, 67, 67
PDS_70	35, 35
G35.20-0.74N	36, 56
LupusIII_19	40, 40
2MASS_J04220069+2657324	40, 48
IM_Lup	41, 41, 43, 46, 46, 48, 53
lkca15	42, 42, 42, 42, 42, 42, 42, 50, 60
LupusIII_1015	42, 66
CY_Tau	43, 43, 61
HD_163296	43, 46, 48, 65, 65
UX_Tau	44, 67
WaOph_6	46, 46
RA16_21_31.920	47, 65
G35.03+0.35	48, 64
CX_Tau	49, 61
GM_Aur	50, 50, 50
VSSG_1	51, 52
HN_Tau	51, 51
V710_Tau	51, 51, 64
HD_100453	52, 66, 66
sz84	54, 54
J04362151+2351165	57, 57
IRAS_04187+1927	58, 58
pds99	66, 66
UX_Tau_A	66, 67

Tabell B.3: DBSCAN utan t-SNE

Objekt	Kluster
TW_Hya	-1, -1, -1, -1, -1, -1, 8, 32, 32
2MASS_J16042165-2130284	-1, -1, -1

Objekt	Kluster
Oph_IRS_48	-1, -1, -1, -1, -1
G35.03+0.35	-1, -1
HD_169142	-1, 19, 19
lkca15	-1, -1, -1, -1, -1, 7, 7, 17, 17
GO_Tau	-1, -1
LupusIII_19	-1, -1
Serpens-FIRS1	-1, -1
RA16_21_31.920	-1, -1
AA_Tau	-1, -1, -1, -1
TCha	-1, -1
RA16_31_31.243	-1, 8
2MASS_J16085553-3902339	-1, 2
HD_34282	-1, -1, -1, -1, -1
CIDA_1	-1, -1, -1
RW_Aur	-1, 1, 1
HD_142527	-1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, 4, 12, 12
DM_Tau	-1, -1, -1, -1, -1, 1, 1, 1
AS_209	-1, -1, 14, 14
HD163296	-1, -1, -1, 25
IM_Lup	-1, -1, -1, 2, 21, 21, 25
UX_Tau_A	-1, -1
IRAS_04187+1927	-1, -1
LupusIII_1015	-1, -1
G35.20-0.74N	-1, -1
LupusIII_21	-1, 1
serp_smm2	-1, 1
UX_Tau	-1, -1
2MASS_J04220069+2657324	-1, -1
Objekt utan namn	-1, -1, 0, 1, 4, 13
pds99	-1, -1
TWA_7	-1, 1
J16042165-2130284	-1, 1
J16090075-1908526	-1, 1
RXJ1842.9-3532	-1, -1, -1, 22, 22
J04295950+2433078	-1, -1
RXJ1852.3-3700	-1, -1, -1, -1, 0
serp_smm9	-1, -1
serp_smm9d	-1, -1
S255_IR	-1, -1
HD_163296	-1, -1, -1, -1, 21
J16082324-1930009	-1, 1
V710_Tau	-1, 1, 1
WaOph_6	-1, 21
LupusIII_91	-1, -1
CX_Tau	-1, 1
IRS_63	-1, 1, 1
RNO_90	-1, 1, 1
GM_Aur	-1, 24, 24

Objekt	Kluster
hd142527	-1, 13
J15583692-2257153	-1, 1
LupusIII_53	-1, 1
HD_100453	1, 3, 3
GSS_31	1, 1
VSSG_1	1, 1
HN_Tau	1, 1
serp_smm4	1, 1
mwc480	1, 1, 1, 1, 1, 1, 1, 1, 8
AS_205	1, 1, 1
RA16_27_13.807	1, 1
serp_smm11	1, 1
CY_Tau	1, 1, 1
Parsamian_21	1, 1
Wa_Oph_6	1, 1
DO_Tau	1, 1
HD102647	1, 1
J16113134-1838259	1, 1
UZ_Tau	1, 1
RA16_32_21.038	1, 2
RY_Tau	1, 1
Hd_142527	6, 6, 13, 26, 26
DM_tau	8, 8
HD_97048	8, 20, 20
V4046_Sgr	9, 9, 23, 23, 27, 27
LkCa_15	10, 10, 28, 28
sz84	15, 15
hd169142	19, 19
PDS_70	29, 29
J04362151+2351165	30, 30

Tabell B.4: DBSCAN med t-SNE

Objekt	Kluster
lkca15	-1, -1, 38, 38, 38, 38, 38, 38, 38
Oph_IRS_48	-1, 5, 65, 106, 106
RA16_31_31.243	-1, 95
RW_Aur	-1, 1, 48
HD_34282	-1, 8, 59, 59, 92
AS_209	-1, 31, 68, 68
TW_Hya	-1, -1, 1, 20, 30, 35, 76, 107, 107
LupusIII_21	-1, -1
AA_Tau	-1, 32, 32, 54
TWA_7	-1, 2
DM_Tau	-1, -1, 1, 31, 33, 42, 64, 64
G35.03+0.35	-1, 7

Objekt	Kluster
serp_smm9d	-1, 66
mwc480	-1, 33, 33, 37, 37, 39, 39, 47, 47
V710_Tau	-1, 10, 10
G35.20-0.74N	-1, 78
LupusIII_53	-1, 23
RXJ1852.3-3700	0, 4, 4, 50, 50
Objekt utan namn	0, 1, 16, 16, 57, 90
GO_Tau	1, 1
CY_Tau	1, 41, 93
Wa_Oph_6	1, 1
S255_IR	1, 28
IRS_63	1, 34, 41
HD163296	1, 1, 1, 71
HD_163296	1, 25, 25, 35, 76
IM_Lup	1, 5, 5, 35, 35, 55, 76
HD_100453	2, 12, 12
AS_205	2, 10, 23
VSSG_1	2, 10
DO_Tau	2, 2
UZ_Tau	2, 47
RY_Tau	2, 2
2MASS_J16085553-3902339	3, 7
RA16_32_21.038	3, 23
2MASS_J16042165-2130284	4, 4, 50
PDS_70	4, 4
RA16_21_31.920	6, 25
J04362151+2351165	7, 7
J16082324-1930009	7, 36
HD_169142	8, 80, 80
J04295950+2433078	8, 98
HN_Tau	10, 10
serp_smm11	10, 65
J16090075-1908526	11, 30
J16042165-2130284	11, 94
RXJ1842.9-3532	12, 12, 12, 12, 57
LupusIII_1015	12, 14
pds99	12, 12
UX_Tau_A	12, 50
HD_142527	16, 16, 16, 16, 16, 16, 16, 16, 16, 26, 61, 61, 61, 61
Hd_142527	16, 26, 26, 97, 97
hd142527	16, 16
LupusIII_19	18, 18
2MASS_J04220069+2657324	18, 46
Serpens-FIRS1	19, 19
GSS_31	21, 23
RNO_90	31, 84, 93
GM_Aur	32, 32, 32
serp_smm2	33, 82

B. Klusterinformation

Objekt	Kluster
HD102647	33, 33
J15583692-2257153	33, 94
serp_smm4	34, 36
WaOph_6	35, 35
DM_tau	39, 39
IRAS_04187+1927	39, 85
CX_Tau	41, 62
RA16_27_13.807	41, 70
TCha	42, 59
Parsamian_21	48, 84
LkCa_15	50, 50, 50, 50
UX_Tau	50, 103
V4046_Sgr	50, 50, 51, 51, 87, 87
J16113134-1838259	56, 56
CIDA_1	57, 89, 89
sz84	72, 72
hd169142	80, 80
LupusIII_91	82, 101
HD_97048	83, 83, 95
serp_smm9	99, 102

C

Kod

All kod inklusive testfilen och en textfil med länk till samtliga observationer som använts i projektet finns tillgängliga i github med namnet “Image_Clustering_Of_Protoplanetary_disks_ALMA”. Länkarna till observationerna ligger i filen som heter “data_links.txt”. Länken till github finns tillgänglig nedan.

https://github.com/theagranstrom/Image_Clustering_Of_Protoplanetary_disks_ALMA

D

Bildmanipuleringsinformation

En utskrift från vår testfil "*test_star_in_same_clusters.py*" på vilka kluster som alla objekt tillhör, samt de rotationsfaktorer som gjorts på bilden och hur stor den var innan zoom. Detta exempel är från vår datamängd som först filtrerats med sigma satt till 7 för filtering och därefter sigma 5 för brusreducering. Därefter DCT, t-SNE och *k*-means med *k* satt till 70 applicerat.

RA16_25_46.630 i kluster:

- Kluster: 0 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [13, 19]

J11122772-7644223 i kluster:

- Kluster: 0 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [16, 14]

J11105359-7725004 i kluster:

- Kluster: 0 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [19, 13]

mwc480 i kluster:

- Kluster: 0 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [26, 21]

- Kluster: 0 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [34, 27]

- Kluster: 6 , rotationsfaktor: 1.8 , storlek innan zoom: [34, 27]

- Kluster: 10 , rotationsfaktor: 1.0 , storlek innan zoom: [23, 19]

- Kluster: 10 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [29, 20]

- Kluster: 12 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [27, 21]

- Kluster: 17 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [36, 23]

- Kluster: 28 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [16, 15]

- Kluster: 58 , rotationsfaktor: 1.5 , storlek innan zoom: [27, 21]

J11092379-7623207 i kluster:

- Kluster: 0 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [16, 14]

GSS_31 i kluster:

- Kluster: 1 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 20]

- Kluster: 19 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 21]

FN_Tau i kluster:

- Kluster: 1 , rotationsfaktor: 1.0 , storlek innan zoom: [19, 20]

SR_4 i kluster:

- Kluster: 1 , rotationsfaktor: 1.0 , storlek innan zoom: [26, 26]

DoAr_24E i kluster:

- Kluster: 1 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 16]

AS_205 i kluster:

- Kluster: 1 , rotationsfaktor: 1.0 , storlek innan zoom: [29, 33]

- Kluster: 51 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [23, 29]

- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [26, 28]

RW_Aur i kluster:

- Kluster: 1 , rotationsfaktor: 1.0 , storlek innan zoom: [26, 19]

D. Bildmanipuleringsinformation

- Kluster: 1 , rotationsfaktor: 1.0 , storlek innan zoom: [27, 21]
- Kluster: 56 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [48, 34]
- Parsamian_21 i kluster:
 - Kluster: 1 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [33, 29]
 - Kluster: 19 , rotationsfaktor: 1.5 , storlek innan zoom: [47, 41]
- DoAr_33 i kluster:
 - Kluster: 1 , rotationsfaktor: 1.0 , storlek innan zoom: [20, 26]
- HT_Lup i kluster:
 - Kluster: 1 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [21, 22]
- WSB_52 i kluster:
 - Kluster: 1 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [22, 23]
- HD_169142 i kluster:
 - Kluster: 2 , rotationsfaktor: 1.0 , storlek innan zoom: [21, 26]
 - Kluster: 27 , rotationsfaktor: 1.0 , storlek innan zoom: [106, 107]
 - Kluster: 27 , rotationsfaktor: 1.0 , storlek innan zoom: [79, 79]
- sigOri_859 i kluster:
 - Kluster: 2 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [13, 13]
- HD_34282 i kluster:
 - Kluster: 2 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [15, 22]
 - Kluster: 3 , rotationsfaktor: 1.5 , storlek innan zoom: [51, 72]
 - Kluster: 3 , rotationsfaktor: 1.5 , storlek innan zoom: [37, 50]
 - Kluster: 3 , rotationsfaktor: 1.5 , storlek innan zoom: [42, 56]
 - Kluster: 15 , rotationsfaktor: 1.6 , storlek innan zoom: [21, 26]
- J11095873-7737088 i kluster:
 - Kluster: 2 , rotationsfaktor: 1.4 , storlek innan zoom: [15, 12]
- RA16_26_03.012 i kluster:
 - Kluster: 2 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [9, 12]
- 2MASS_J11075809-7742413 i kluster:
 - Kluster: 2 , rotationsfaktor: 0.625 , storlek innan zoom: [13, 14]
- sigOri_774 i kluster:
 - Kluster: 2 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [21, 28]
- J11120984-7634366 i kluster:
 - Kluster: 2 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 12]
- J04295950+2433078 i kluster:
 - Kluster: 2 , rotationsfaktor: 0.5555555555555556 , storlek innan zoom: [16, 16]
 - Kluster: 60 , rotationsfaktor: 0.625 , storlek innan zoom: [16, 19]
- RA16_26_48.638 i kluster:
 - Kluster: 2 , rotationsfaktor: 1.0 , storlek innan zoom: [9, 13]
- TCha i kluster:
 - Kluster: 3 , rotationsfaktor: 2.5 , storlek innan zoom: [72, 126]
 - Kluster: 16 , rotationsfaktor: 2.5 , storlek innan zoom: [55, 86]
- RA16_27_39.811 i kluster:
 - Kluster: 4 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]
- RA16_27_04.097 i kluster:
 - Kluster: 4 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [13, 16]
- RA16_31_33.444 i kluster:
 - Kluster: 4 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 19]
- RA16_27_02.993 i kluster:
 - Kluster: 4 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]

RA16_27_06.754 i kluster:

- Kluster: 4 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]

AA_Tau i kluster:

- Kluster: 4 , rotationsfaktor: 1.0 , storlek innan zoom: [29, 44]
- Kluster: 32 , rotationsfaktor: 1.0 , storlek innan zoom: [51, 92]
- Kluster: 50 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [42, 70]
- Kluster: 50 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [33, 51]

RA16_27_03.578 i kluster:

- Kluster: 4 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

RA16_28_23.323 i kluster:

- Kluster: 4 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]

serp_smm9 i kluster:

- Kluster: 4 , rotationsfaktor: 0.625 , storlek innan zoom: [27, 23]
- Kluster: 19 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [34, 22]

RA16_26_10.320 i kluster:

- Kluster: 4 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

Sz_65 i kluster:

- Kluster: 4 , rotationsfaktor: 0.875 , storlek innan zoom: [20, 23]

i kluster:

- Kluster: 5 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [90, 96]
- Kluster: 5 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [54, 69]
- Kluster: 15 , rotationsfaktor: 0.875 , storlek innan zoom: [41, 37]
- Kluster: 24 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [23, 20]
- Kluster: 56 , rotationsfaktor: 1.0 , storlek innan zoom: [43, 43]
- Kluster: 66 , rotationsfaktor: 1.0 , storlek innan zoom: [40, 40]

HD_142527 i kluster:

- Kluster: 5 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [77, 90]
- Kluster: 5 , rotationsfaktor: 1.0 , storlek innan zoom: [89, 106]
- Kluster: 5 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [62, 65]
- Kluster: 5 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [54, 63]
- Kluster: 5 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [65, 76]
- Kluster: 5 , rotationsfaktor: 1.0 , storlek innan zoom: [89, 110]
- Kluster: 5 , rotationsfaktor: 1.0 , storlek innan zoom: [57, 63]
- Kluster: 5 , rotationsfaktor: 1.0 , storlek innan zoom: [64, 72]
- Kluster: 5 , rotationsfaktor: 1.0 , storlek innan zoom: [55, 70]
- Kluster: 45 , rotationsfaktor: 1 , storlek innan zoom: [27, 43]
- Kluster: 45 , rotationsfaktor: 1.0 , storlek innan zoom: [20, 28]
- Kluster: 45 , rotationsfaktor: 1 , storlek innan zoom: [49, 79]
- Kluster: 45 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [36, 58]
- Kluster: 45 , rotationsfaktor: 1 , storlek innan zoom: [40, 43]

Hd_142527 i kluster:

- Kluster: 5 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [54, 65]
- Kluster: 28 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [19, 21]
- Kluster: 28 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [19, 21]
- Kluster: 45 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [35, 62]
- Kluster: 45 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [35, 61]

hd142527 i kluster:

- Kluster: 5 , rotationsfaktor: 1.0 , storlek innan zoom: [43, 51]
- Kluster: 5 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [48, 54]

D. Bildmanipuleringsinformation

DM_tau i kluster:

- Kluster: 6 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [15, 19]
- Kluster: 6 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [15, 19]

MYLup i kluster:

- Kluster: 6 , rotationsfaktor: 0.4 , storlek innan zoom: [40, 56]

J11142454-7733062 i kluster:

- Kluster: 6 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 13]

J11173700-7704381 i kluster:

- Kluster: 6 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [15, 13]

Serpens-FIRS1 i kluster:

- Kluster: 7 , rotationsfaktor: 1.0 , storlek innan zoom: [47, 43]
- Kluster: 7 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [75, 69]

2MASS_J16020287-2236139 i kluster:

- Kluster: 8 , rotationsfaktor: 1.4 , storlek innan zoom: [12, 13]

serp_smm2 i kluster:

- Kluster: 8 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [12, 13]
- Kluster: 12 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [13, 15]

J11040425-7639328 i kluster:

- Kluster: 8 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 8]

J16014086-2258103 i kluster:

- Kluster: 8 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [12, 14]

2MASS_J11071206-7632232 i kluster:

- Kluster: 8 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [13, 15]

RA16_27_37.229 i kluster:

- Kluster: 8 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [9, 14]

LupusIII_91 i kluster:

- Kluster: 8 , rotationsfaktor: 1.2 , storlek innan zoom: [12, 9]
- Kluster: 47 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 12]

RA16_31_52.452 i kluster:

- Kluster: 8 , rotationsfaktor: 1.0 , storlek innan zoom: [9, 13]

2M0444 i kluster:

- Kluster: 9 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [27, 21]

2MASS_J16075796-2040087 i kluster:

- Kluster: 9 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 15]

2MASS_J16024152-2138245 i kluster:

- Kluster: 9 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 15]

AS_209 i kluster:

- Kluster: 9 , rotationsfaktor: 1.0 , storlek innan zoom: [35, 41]
- Kluster: 9 , rotationsfaktor: 1.0 , storlek innan zoom: [33, 40]
- Kluster: 9 , rotationsfaktor: 1.0 , storlek innan zoom: [35, 40]
- Kluster: 29 , rotationsfaktor: 1.0 , storlek innan zoom: [47, 55]

RNO_90 i kluster:

- Kluster: 9 , rotationsfaktor: 1.0 , storlek innan zoom: [26, 23]
- Kluster: 19 , rotationsfaktor: 1.0 , storlek innan zoom: [23, 22]
- Kluster: 43 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [42, 28]

Hn_22 i kluster:

- Kluster: 9 , rotationsfaktor: 1.0 , storlek innan zoom: [20, 29]

DM_Tau i kluster:

- Kluster: 9 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [27, 23]

- Kluster: 12 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [23, 21]
- Kluster: 16 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [36, 33]
- Kluster: 19 , rotationsfaktor: 1.0 , storlek innan zoom: [43, 40]
- Kluster: 30 , rotationsfaktor: 1.0 , storlek innan zoom: [26, 22]
- Kluster: 44 , rotationsfaktor: 1.0 , storlek innan zoom: [44, 37]
- Kluster: 44 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [47, 40]
- Kluster: 56 , rotationsfaktor: 1.0 , storlek innan zoom: [29, 26]
- LupusIII_1014 i kluster:
 - Kluster: 9 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [30, 28]
- DO_Tau i kluster:
 - Kluster: 9 , rotationsfaktor: 0.875 , storlek innan zoom: [30, 23]
 - Kluster: 59 , rotationsfaktor: 0.875 , storlek innan zoom: [35, 27]
- MY_Lup i kluster:
 - Kluster: 10 , rotationsfaktor: 0.5555555555555556 , storlek innan zoom: [27, 35]
- RA16_25_56.148 i kluster:
 - Kluster: 10 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]
- J11062554-7633418 i kluster:
 - Kluster: 10 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [21, 15]
- RA16_26_58.423 i kluster:
 - Kluster: 10 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 19]
- UZ_Tau i kluster:
 - Kluster: 10 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 21]
 - Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [20, 21]
- J11004022-7619280 i kluster:
 - Kluster: 10 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [22, 19]
- RY_Tau i kluster:
 - Kluster: 10 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [44, 28]
 - Kluster: 52 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [40, 26]
- TW_Hya i kluster:
 - Kluster: 11 , rotationsfaktor: 0.75 , storlek innan zoom: [219, 327]
 - Kluster: 13 , rotationsfaktor: 0.875 , storlek innan zoom: [57, 82]
 - Kluster: 22 , rotationsfaktor: 1.0 , storlek innan zoom: [21, 26]
 - Kluster: 25 , rotationsfaktor: 1.0 , storlek innan zoom: [43, 37]
 - Kluster: 46 , rotationsfaktor: 1.0 , storlek innan zoom: [34, 33]
 - Kluster: 48 , rotationsfaktor: 1.8 , storlek innan zoom: [245, 405]
 - Kluster: 48 , rotationsfaktor: 1.0 , storlek innan zoom: [42, 41]
 - Kluster: 48 , rotationsfaktor: 1.0 , storlek innan zoom: [36, 37]
 - Kluster: 53 , rotationsfaktor: 0.875 , storlek innan zoom: [41, 41]
- HD163296 i kluster:
 - Kluster: 11 , rotationsfaktor: 1.5 , storlek innan zoom: [79, 78]
 - Kluster: 18 , rotationsfaktor: 1.5 , storlek innan zoom: [47, 49]
 - Kluster: 43 , rotationsfaktor: 1.5 , storlek innan zoom: [37, 42]
 - Kluster: 43 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [35, 35]
- sz98 i kluster:
 - Kluster: 11 , rotationsfaktor: 1.2857142857142858 , storlek innan zoom: [91, 134]
- V1094_Sco i kluster:
 - Kluster: 12 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [23, 30]
- J11104959-7717517 i kluster:
 - Kluster: 12 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 14]

D. Bildmanipuleringsinformation

HD102647 i kluster:

- Kluster: 12 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [15, 14]
- Kluster: 12 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [19, 16]

HD104237 i kluster:

- Kluster: 12 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [20, 15]

2MASS_J11111083-7641574 i kluster:

- Kluster: 12 , rotationsfaktor: 2.25 , storlek innan zoom: [33, 23]

LupusIV_153 i kluster:

- Kluster: 12 , rotationsfaktor: 1.8 , storlek innan zoom: [30, 22]

J11083905-7716042 i kluster:

- Kluster: 13 , rotationsfaktor: 1.4 , storlek innan zoom: [13, 9]

2MASS_J16085553-3902339 i kluster:

- Kluster: 13 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [9, 9]
- Kluster: 38 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [14, 15]

JH_223 i kluster:

- Kluster: 13 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [13, 13]

WSB_4 i kluster:

- Kluster: 13 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [9, 9]

TWA_34 i kluster:

- Kluster: 13 , rotationsfaktor: 1.0 , storlek innan zoom: [8, 14]

J11120351-7726009 i kluster:

- Kluster: 13 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [9, 8]

J16082324-1930009 i kluster:

- Kluster: 13 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [9, 12]
- Kluster: 28 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 19]

2MASS_J11085464-7702129 i kluster:

- Kluster: 13 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [13, 12]

Oph_IRS_48 i kluster:

- Kluster: 14 , rotationsfaktor: 1.8 , storlek innan zoom: [23, 44]
- Kluster: 19 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [21, 30]
- Kluster: 37 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [26, 43]
- Kluster: 37 , rotationsfaktor: 1.6 , storlek innan zoom: [28, 48]
- Kluster: 41 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 19]

RA16_27_13.714 i kluster:

- Kluster: 14 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 15]

J16024152-2138245 i kluster:

- Kluster: 14 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]

J16181904-2028479 i kluster:

- Kluster: 14 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 16]

J16035767-2031055 i kluster:

- Kluster: 14 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [13, 15]

J16135434-2320342 i kluster:

- Kluster: 14 , rotationsfaktor: 0.75 , storlek innan zoom: [13, 16]

EPIC_204107757 i kluster:

- Kluster: 14 , rotationsfaktor: 1.0 , storlek innan zoom: [21, 23]

FO_Tau i kluster:

- Kluster: 14 , rotationsfaktor: 0.75 , storlek innan zoom: [19, 21]

CIDA_1 i kluster:

- Kluster: 15 , rotationsfaktor: 1.0 , storlek innan zoom: [23, 22]

- Kluster: 44 , rotationsfaktor: 0.875 , storlek innan zoom: [41, 29]
- Kluster: 44 , rotationsfaktor: 0.875 , storlek innan zoom: [36, 27]
- V4046_Sgr i kluster:
 - Kluster: 15 , rotationsfaktor: 1.0 , storlek innan zoom: [26, 27]
 - Kluster: 15 , rotationsfaktor: 1.0 , storlek innan zoom: [33, 35]
 - Kluster: 16 , rotationsfaktor: 0.875 , storlek innan zoom: [55, 65]
 - Kluster: 16 , rotationsfaktor: 0.875 , storlek innan zoom: [54, 63]
 - Kluster: 63 , rotationsfaktor: 0.875 , storlek innan zoom: [48, 48]
 - Kluster: 63 , rotationsfaktor: 0.875 , storlek innan zoom: [40, 37]
- LupusIII_1010 i kluster:
 - Kluster: 15 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [23, 21]
- 2MASS_J15583692-2257153 i kluster:
 - Kluster: 15 , rotationsfaktor: 1.0 , storlek innan zoom: [29, 28]
- Sz_58 i kluster:
 - Kluster: 15 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [29, 43]
- RXJ1842.9-3532 i kluster:
 - Kluster: 15 , rotationsfaktor: 0.75 , storlek innan zoom: [28, 27]
 - Kluster: 20 , rotationsfaktor: 0.875 , storlek innan zoom: [51, 47]
 - Kluster: 20 , rotationsfaktor: 0.875 , storlek innan zoom: [41, 41]
 - Kluster: 20 , rotationsfaktor: 0.875 , storlek innan zoom: [42, 41]
 - Kluster: 20 , rotationsfaktor: 0.875 , storlek innan zoom: [51, 48]
- LkCa_15 i kluster:
 - Kluster: 16 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [48, 55]
 - Kluster: 16 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [62, 75]
 - Kluster: 16 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [50, 61]
 - Kluster: 16 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [55, 62]
- HD_100546 i kluster:
 - Kluster: 16 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [86, 83]
- mwc_758 i kluster:
 - Kluster: 16 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [44, 40]
- DE_Tau i kluster:
 - Kluster: 17 , rotationsfaktor: 1.6 , storlek innan zoom: [33, 28]
- IRAS13481-6124 i kluster:
 - Kluster: 17 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [54, 48]
- J11183572-7935548 i kluster:
 - Kluster: 17 , rotationsfaktor: 1.0 , storlek innan zoom: [19, 15]
- RA16_31_35.647 i kluster:
 - Kluster: 17 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]
- 2MASS_J16232807-4015368 i kluster:
 - Kluster: 17 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [15, 16]
- J11040909-7627193 i kluster:
 - Kluster: 17 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [19, 14]
- J16123916-1859284 i kluster:
 - Kluster: 17 , rotationsfaktor: 0.75 , storlek innan zoom: [14, 20]
- 2MASS_J15530132-2114135 i kluster:
 - Kluster: 17 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 15]
- Sz_63 i kluster:
 - Kluster: 17 , rotationsfaktor: 0.75 , storlek innan zoom: [19, 34]
- J10590699-7701404 i kluster:

- Kluster: 17 , rotationsfaktor: 1.0 , storlek innan zoom: [20, 16]
- DQ_Tau i kluster:
- Kluster: 17 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 14]
- GO_Tau i kluster:
- Kluster: 18 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [56, 42]
- Kluster: 18 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [50, 37]
- HL_Tau i kluster:
- Kluster: 18 , rotationsfaktor: 1.5 , storlek innan zoom: [124, 119]
- HT_Lupi i kluster:
- Kluster: 18 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [82, 77]
- Wa_Oph_6 i kluster:
- Kluster: 18 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [42, 30]
- Kluster: 18 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [42, 30]
- S255_IR i kluster:
- Kluster: 18 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [90, 70]
- Kluster: 33 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [21, 21]
- IRS_63 i kluster:
- Kluster: 18 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [275, 222]
- Kluster: 23 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [23, 23]
- Kluster: 59 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [34, 34]
- G023.01-00.41 i kluster:
- Kluster: 18 , rotationsfaktor: 3.3333333333333335 , storlek innan zoom: [100, 78]
- 2MASS_J04141188+2811535 i kluster:
- Kluster: 19 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [26, 23]
- serp_smm11 i kluster:
- Kluster: 19 , rotationsfaktor: 1.0 , storlek innan zoom: [19, 20]
- Kluster: 51 , rotationsfaktor: 0.75 , storlek innan zoom: [16, 21]
- FP_Tau i kluster:
- Kluster: 19 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [19, 21]
- KK_Oph i kluster:
- Kluster: 19 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [15, 15]
- Haro_6-37 i kluster:
- Kluster: 19 , rotationsfaktor: 0.75 , storlek innan zoom: [40, 30]
- LupusIII_53 i kluster:
- Kluster: 19 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [20, 19]
- Kluster: 30 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [15, 19]
- AK_Sco i kluster:
- Kluster: 20 , rotationsfaktor: 0.5555555555555556 , storlek innan zoom: [47, 56]
- HD_143006 i kluster:
- Kluster: 20 , rotationsfaktor: 1.0 , storlek innan zoom: [37, 36]
- CQ_Tau i kluster:
- Kluster: 20 , rotationsfaktor: 0.875 , storlek innan zoom: [34, 35]
- V892_Tau i kluster:
- Kluster: 20 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [37, 36]
- v4046_sgr i kluster:
- Kluster: 20 , rotationsfaktor: 1.0 , storlek innan zoom: [28, 34]
- J10555973-7724399 i kluster:
- Kluster: 21 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [16, 13]
- J11072074-7738073 i kluster:

- Kluster: 21 , rotationsfaktor: 1.4 , storlek innan zoom: [15, 12]
RA16_31_31.243 i kluster:
- Kluster: 21 , rotationsfaktor: 0.3 , storlek innan zoom: [56, 65]
- Kluster: 22 , rotationsfaktor: 0.625 , storlek innan zoom: [15, 19]
J11105333-7634319 i kluster:
- Kluster: 21 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [16, 13]
sigOri_1036 i kluster:
- Kluster: 21 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [15, 15]
J11085367-7521359 i kluster:
- Kluster: 21 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 13]
J11094742-7726290 i kluster:
- Kluster: 21 , rotationsfaktor: 0.875 , storlek innan zoom: [21, 16]
Sz_59 i kluster:
- Kluster: 22 , rotationsfaktor: 0.875 , storlek innan zoom: [14, 23]
J16111330-2019029 i kluster:
- Kluster: 22 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [14, 19]
RA16_26_46.418 i kluster:
- Kluster: 22 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [12, 14]
J1653456-2242421 i kluster:
- Kluster: 22 , rotationsfaktor: 0.75 , storlek innan zoom: [13, 16]
TWA_32 i kluster:
- Kluster: 22 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [14, 19]
J11022491-7733357 i kluster:
- Kluster: 22 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [19, 15]
RA16_25_36.722 i kluster:
- Kluster: 22 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [12, 15]
RA16_26_45.014 i kluster:
- Kluster: 22 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [15, 20]
HD_97048 i kluster:
- Kluster: 22 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [30, 23]
- Kluster: 68 , rotationsfaktor: 1.0 , storlek innan zoom: [97, 72]
- Kluster: 68 , rotationsfaktor: 1.0 , storlek innan zoom: [104, 78]
RA16_27_21.454 i kluster:
- Kluster: 22 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [12, 15]
SSTc2d_J130059.3-771403 i kluster:
- Kluster: 22 , rotationsfaktor: 0.875 , storlek innan zoom: [16, 22]
RA16_31_36.770 i kluster:
- Kluster: 22 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]
LupusIII_1013 i kluster:
- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 14]
RA16_26_17.227 i kluster:
- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]
IRAS12496-7650 i kluster:
- Kluster: 23 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [27, 36]
serp_smm4 i kluster:
- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [19, 21]
- Kluster: 28 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [22, 22]
HD100546 i kluster:
- Kluster: 23 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [19, 15]

GQ_Lup i kluster:

- Kluster: 23 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [20, 23]

RA16_26_18.862 i kluster:

- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 16]

Sz_50 i kluster:

- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [19, 29]

RA16_27_38.309 i kluster:

- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]

RA16_27_19.198 i kluster:

- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

LupusIII_21 i kluster:

- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 16]

- Kluster: 37 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [19, 16]

RA16_26_30.446 i kluster:

- Kluster: 23 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

J16054540-2023088 i kluster:

- Kluster: 24 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [13, 16]

Hn_25 i kluster:

- Kluster: 24 , rotationsfaktor: 0.75 , storlek innan zoom: [16, 22]

RA16_23_34.814 i kluster:

- Kluster: 24 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [9, 12]

RA16_27_30.168 i kluster:

- Kluster: 24 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [12, 14]

Sz_53 i kluster:

- Kluster: 24 , rotationsfaktor: 0.75 , storlek innan zoom: [15, 26]

J16062196-1928445 i kluster:

- Kluster: 24 , rotationsfaktor: 0.75 , storlek innan zoom: [13, 16]

J16042165-2130284 i kluster:

- Kluster: 24 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 15]

- Kluster: 34 , rotationsfaktor: 0.75 , storlek innan zoom: [16, 21]

EM_star_AS_209 i kluster:

- Kluster: 24 , rotationsfaktor: 0.875 , storlek innan zoom: [47, 57]

RA16_27_40.267 i kluster:

- Kluster: 24 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 14]

J16072625-2432079 i kluster:

- Kluster: 24 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 16]

2MASS_J04230607+2801194 i kluster:

- Kluster: 24 , rotationsfaktor: 1.4 , storlek innan zoom: [19, 14]

J15583692-2257153 i kluster:

- Kluster: 24 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 14]

- Kluster: 28 , rotationsfaktor: 0.75 , storlek innan zoom: [15, 20]

RA16_27_55.567 i kluster:

- Kluster: 24 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 19]

RA16_32_21.924 i kluster:

- Kluster: 25 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [9, 9]

J11100469-7635452 i kluster:

- Kluster: 25 , rotationsfaktor: 1.2 , storlek innan zoom: [9, 8]

2MASS_J11091812-7630292 i kluster:

- Kluster: 25 , rotationsfaktor: 0.5 , storlek innan zoom: [9, 9]

J10580597-7711501 i kluster:

- Kluster: 25 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [9, 8]

DP_Tau i kluster:

- Kluster: 25 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 14]

RA16_28_57.854 i kluster:

- Kluster: 25 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [8, 9]

2MASS_J15534211-2049282 i kluster:

- Kluster: 25 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [8, 13]

J16090075-1908526 i kluster:

- Kluster: 25 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [8, 13]

- Kluster: 34 , rotationsfaktor: 0.75 , storlek innan zoom: [14, 19]

2MASS_J16063539-2516510 i kluster:

- Kluster: 25 , rotationsfaktor: 1.0 , storlek innan zoom: [8, 8]

J11100369-7633291 i kluster:

- Kluster: 25 , rotationsfaktor: 1.4 , storlek innan zoom: [9, 8]

IS_Tau i kluster:

- Kluster: 25 , rotationsfaktor: 8.0 , storlek innan zoom: [15, 15]

J11065939-7530559 i kluster:

- Kluster: 25 , rotationsfaktor: 0.5714285714285714 , storlek innan zoom: [13, 9]

J12074836-3900043 i kluster:

- Kluster: 25 , rotationsfaktor: 0.8888888888888888 , storlek innan zoom: [9, 9]

RA16_27_39.010 i kluster:

- Kluster: 26 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]

Hn_23 i kluster:

- Kluster: 26 , rotationsfaktor: 0.75 , storlek innan zoom: [20, 29]

V346_Nor i kluster:

- Kluster: 26 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [15, 16]

RA16_26_56.765 i kluster:

- Kluster: 26 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]

RA16_28_45.262 i kluster:

- Kluster: 26 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

RA16_27_11.167 i kluster:

- Kluster: 26 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

RA16_27_38.928 i kluster:

- Kluster: 26 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

RA16_21_45.127 i kluster:

- Kluster: 26 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 16]

RA16_26_18.977 i kluster:

- Kluster: 26 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]

RA16_31_52.109 i kluster:

- Kluster: 26 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]

hd169142 i kluster:

- Kluster: 27 , rotationsfaktor: 1.0 , storlek innan zoom: [70, 70]

- Kluster: 27 , rotationsfaktor: 1.0 , storlek innan zoom: [83, 82]

HD169142 i kluster:

- Kluster: 27 , rotationsfaktor: 1.0 , storlek innan zoom: [103, 103]

WSB_35 i kluster:

- Kluster: 28 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 15]

RA16_26_24.036 i kluster:

D. Bildmanipuleringsinformation

- Kluster: 28 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 19]
RA16_22_18.521 i kluster:
- Kluster: 28 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 15]
RA16_27_17.580 i kluster:
- Kluster: 28 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]
J10590108-7722407 i kluster:
- Kluster: 28 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [19, 15]
RA16_28_54.058 i kluster:
- Kluster: 28 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 15]
2MASS_J16111330-2019029 i kluster:
- Kluster: 28 , rotationsfaktor: 1.4 , storlek innan zoom: [13, 14]
IT_Tau i kluster:
- Kluster: 29 , rotationsfaktor: 1.0 , storlek innan zoom: [22, 16]
LupusI_14 i kluster:
- Kluster: 29 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 20]
LupusI_15 i kluster:
- Kluster: 29 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 20]
WSB_71 i kluster:
- Kluster: 29 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [19, 19]
SU_Aur i kluster:
- Kluster: 29 , rotationsfaktor: 1.0 , storlek innan zoom: [28, 22]
Sz_129 i kluster:
- Kluster: 29 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [33, 30]
Sz_61 i kluster:
- Kluster: 29 , rotationsfaktor: 1.0 , storlek innan zoom: [27, 37]
RA16_27_13.807 i kluster:
- Kluster: 29 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [26, 28]
- Kluster: 62 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]
2MASS_J16153456-2242421 i kluster:
- Kluster: 29 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 15]
HD144668 i kluster:
- Kluster: 30 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 15]
V807_Tau i kluster:
- Kluster: 30 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [22, 19]
LupusIII_137 i kluster:
- Kluster: 30 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 16]
DR_Tau i kluster:
- Kluster: 30 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [28, 27]
EX_Lup i kluster:
- Kluster: 30 , rotationsfaktor: 0.875 , storlek innan zoom: [22, 26]
RA16_32_21.038 i kluster:
- Kluster: 30 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [22, 27]
- Kluster: 38 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]
PBB2002_UScoJ160900.7-190852 i kluster:
- Kluster: 30 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 20]
Elias_20 i kluster:
- Kluster: 30 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [30, 28]
2MASS_J16001844-2230114 i kluster:
- Kluster: 31 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 14]

J16154416-1921171 i kluster:

- Kluster: 31 , rotationsfaktor: 0.75 , storlek innan zoom: [14, 19]

RA16_31_44.568 i kluster:

- Kluster: 31 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]

WSB_18 i kluster:

- Kluster: 31 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [13, 14]

Sz_49 i kluster:

- Kluster: 31 , rotationsfaktor: 1.0 , storlek innan zoom: [23, 28]

2MASS_J16164198-3650456 i kluster:

- Kluster: 31 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 16]

J11095407-7629253 i kluster:

- Kluster: 31 , rotationsfaktor: 1.4 , storlek innan zoom: [19, 13]

UY_Aur i kluster:

- Kluster: 32 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [40, 27]

HD36112 i kluster:

- Kluster: 32 , rotationsfaktor: 1.0 , storlek innan zoom: [34, 29]

2M1207 i kluster:

- Kluster: 33 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 13]

GH_Tau i kluster:

- Kluster: 33 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [20, 15]

2MASS_J16154416-1921171 i kluster:

- Kluster: 33 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 14]

TWA_7 i kluster:

- Kluster: 33 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [22, 29]

- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 20]

serp_smm9d i kluster:

- Kluster: 33 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 28]

- Kluster: 55 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [19, 34]

RA16_26_23.671 i kluster:

- Kluster: 34 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 20]

J16075796-2040087 i kluster:

- Kluster: 34 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 19]

HD97048 i kluster:

- Kluster: 34 , rotationsfaktor: 1.0 , storlek innan zoom: [33, 26]

J11100010-7634578 i kluster:

- Kluster: 34 , rotationsfaktor: 1.0 , storlek innan zoom: [20, 16]

RA16_26_24.067 i kluster:

- Kluster: 34 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 19]

J11101141-7635292 i kluster:

- Kluster: 34 , rotationsfaktor: 1.4 , storlek innan zoom: [19, 14]

J16113134-1838259 i kluster:

- Kluster: 34 , rotationsfaktor: 0.75 , storlek innan zoom: [15, 20]

- Kluster: 34 , rotationsfaktor: 0.75 , storlek innan zoom: [15, 20]

RYTau i kluster:

- Kluster: 34 , rotationsfaktor: 0.5555555555555556 , storlek innan zoom: [50, 33]

2MASS_J16042165-2130284 i kluster:

- Kluster: 35 , rotationsfaktor: 1.0 , storlek innan zoom: [77, 78]

- Kluster: 35 , rotationsfaktor: 1.0 , storlek innan zoom: [71, 72]

- Kluster: 67 , rotationsfaktor: 1.0 , storlek innan zoom: [33, 34]

rx_j1604.3-2130 i kluster:

- Kluster: 35 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [58, 62]

RXJ1852.3-3700 i kluster:

- Kluster: 35 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [47, 51]

- Kluster: 35 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [47, 51]

- Kluster: 66 , rotationsfaktor: 1.0 , storlek innan zoom: [28, 28]

- Kluster: 67 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [40, 48]

- Kluster: 67 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [37, 44]

PDS_70 i kluster:

- Kluster: 35 , rotationsfaktor: 1.2857142857142858 , storlek innan zoom: [77, 58]

- Kluster: 35 , rotationsfaktor: 1.2857142857142858 , storlek innan zoom: [78, 61]

Sz91 i kluster:

- Kluster: 35 , rotationsfaktor: 0.875 , storlek innan zoom: [76, 56]

2MASS_J16032225-2413111 i kluster:

- Kluster: 36 , rotationsfaktor: 1.0 , storlek innan zoom: [9, 12]

2MASS_J16123916-1859284 i kluster:

- Kluster: 36 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 14]

HO_Lup i kluster:

- Kluster: 36 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

G35.20-0.74N i kluster:

- Kluster: 36 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 16]

- Kluster: 56 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [33, 28]

2MASS_J16030161-2207523 i kluster:

- Kluster: 36 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [13, 12]

HD98922 i kluster:

- Kluster: 37 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 14]

2MASS_J16181904-2028479 i kluster:

- Kluster: 37 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 15]

DN_Tau i kluster:

- Kluster: 37 , rotationsfaktor: 0.875 , storlek innan zoom: [29, 30]

rho_Oph_102 i kluster:

- Kluster: 37 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 21]

GK_Tau i kluster:

- Kluster: 37 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [20, 15]

2MASS_J16020757-2257467 i kluster:

- Kluster: 38 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 15]

J15582981-2310077 i kluster:

- Kluster: 38 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 19]

RA16_26_23.400 i kluster:

- Kluster: 38 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]

RA16_26_27.530 i kluster:

- Kluster: 38 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]

J11074366-7739411 i kluster:

- Kluster: 38 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [19, 14]

RA16_26_36.816 i kluster:

- Kluster: 38 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]

RA16_22_24.950 i kluster:

- Kluster: 38 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]

RA16_31_43.754 i kluster:

-
- Kluster: 38 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 15]
RA16_27_15.439 i kluster:
 - Kluster: 38 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]
J11044258-7741571 i kluster:
 - Kluster: 39 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [14, 12]
SSTtau042021.4+281349 i kluster:
 - Kluster: 39 , rotationsfaktor: 5.0 , storlek innan zoom: [76, 29]
IRAS_04302+2247 i kluster:
 - Kluster: 39 , rotationsfaktor: 1.5 , storlek innan zoom: [64, 16]
2MASS_J04414825+2534304 i kluster:
 - Kluster: 39 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [15, 12]
Sz_46 i kluster:
 - Kluster: 39 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 21]
2MASS_J04161210+2756385 i kluster:
 - Kluster: 39 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [16, 13]
RA16_25_06.905 i kluster:
 - Kluster: 39 , rotationsfaktor: 1.0 , storlek innan zoom: [9, 13]
2MASS_J16072747-2059442 i kluster:
 - Kluster: 39 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [9, 12]
RA16_27_09.067 i kluster:
 - Kluster: 39 , rotationsfaktor: 1.0 , storlek innan zoom: [9, 13]
FS_Tau i kluster:
 - Kluster: 39 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [15, 12]
LupusIII_19 i kluster:
 - Kluster: 40 , rotationsfaktor: 2.25 , storlek innan zoom: [42, 28]
 - Kluster: 40 , rotationsfaktor: 1.8 , storlek innan zoom: [35, 23]
2MASS_J04220069+2657324 i kluster:
 - Kluster: 40 , rotationsfaktor: 2.5 , storlek innan zoom: [36, 27]
 - Kluster: 48 , rotationsfaktor: 4.333333333333333 , storlek innan zoom: [217, 154]
RA16_27_32.122 i kluster:
 - Kluster: 40 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [9, 13]
2MASS_J11160287-7624533 i kluster:
 - Kluster: 40 , rotationsfaktor: 2.25 , storlek innan zoom: [23, 22]
RA16_26_17.057 i kluster:
 - Kluster: 41 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 14]
2MASS_J16014086-2258103 i kluster:
 - Kluster: 41 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [13, 13]
IM_Lup i kluster:
 - Kluster: 41 , rotationsfaktor: 1.0 , storlek innan zoom: [19, 21]
 - Kluster: 41 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [19, 27]
 - Kluster: 43 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [27, 26]
 - Kluster: 46 , rotationsfaktor: 1.2857142857142858 , storlek innan zoom: [29, 27]
 - Kluster: 46 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [28, 27]
 - Kluster: 48 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [23, 21]
 - Kluster: 53 , rotationsfaktor: 1.0 , storlek innan zoom: [19, 20]
2MASS_J16064102-2455489 i kluster:
 - Kluster: 41 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 13]
XZ_Tau i kluster:
 - Kluster: 42 , rotationsfaktor: 2.25 , storlek innan zoom: [22, 23]

lkca15 i kluster:

- Kluster: 42 , rotationsfaktor: 0.875 , storlek innan zoom: [35, 40]
- Kluster: 42 , rotationsfaktor: 0.75 , storlek innan zoom: [28, 30]
- Kluster: 42 , rotationsfaktor: 0.875 , storlek innan zoom: [36, 37]
- Kluster: 42 , rotationsfaktor: 0.875 , storlek innan zoom: [34, 35]
- Kluster: 42 , rotationsfaktor: 0.875 , storlek innan zoom: [29, 30]
- Kluster: 42 , rotationsfaktor: 0.875 , storlek innan zoom: [27, 29]
- Kluster: 42 , rotationsfaktor: 0.75 , storlek innan zoom: [23, 27]
- Kluster: 50 , rotationsfaktor: 1.0 , storlek innan zoom: [27, 29]
- Kluster: 60 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 19]

RA16_25_02.083 i kluster:

- Kluster: 42 , rotationsfaktor: 0.75 , storlek innan zoom: [13, 23]

LupusIII_50 i kluster:

- Kluster: 42 , rotationsfaktor: 1.8 , storlek innan zoom: [21, 35]

LupusIII_1015 i kluster:

- Kluster: 42 , rotationsfaktor: 1.8 , storlek innan zoom: [23, 40]
- Kluster: 66 , rotationsfaktor: 1.5 , storlek innan zoom: [29, 48]

DL_Tau i kluster:

- Kluster: 43 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [50, 55]

CY_Tau i kluster:

- Kluster: 43 , rotationsfaktor: 1.0 , storlek innan zoom: [35, 34]
- Kluster: 43 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [48, 43]
- Kluster: 61 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [43, 37]

J16011549-4152351 i kluster:

- Kluster: 43 , rotationsfaktor: 2.5 , storlek innan zoom: [86, 61]

CI_Tau i kluster:

- Kluster: 43 , rotationsfaktor: 0.875 , storlek innan zoom: [54, 37]

HD_163296 i kluster:

- Kluster: 43 , rotationsfaktor: 1.2857142857142858 , storlek innan zoom: [55, 62]
- Kluster: 46 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [28, 29]
- Kluster: 48 , rotationsfaktor: 1.5 , storlek innan zoom: [51, 54]
- Kluster: 65 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [16, 20]
- Kluster: 65 , rotationsfaktor: 1.0 , storlek innan zoom: [20, 22]

UX_Tau i kluster:

- Kluster: 44 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [37, 35]
- Kluster: 67 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [35, 29]

MHO_1 i kluster:

- Kluster: 46 , rotationsfaktor: 0.875 , storlek innan zoom: [30, 29]

WaOph_6 i kluster:

- Kluster: 46 , rotationsfaktor: 1.0 , storlek innan zoom: [30, 27]
- Kluster: 46 , rotationsfaktor: 1.0 , storlek innan zoom: [29, 28]

DoAr_25 i kluster:

- Kluster: 46 , rotationsfaktor: 2.0 , storlek innan zoom: [35, 62]

J11081509-7733531 i kluster:

- Kluster: 47 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [20, 15]

RA16_27_15.866 i kluster:

- Kluster: 47 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 12]

Sz_81 i kluster:

- Kluster: 47 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 19]

RA16_21_31.920 i kluster:

- Kluster: 47 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [20, 16]

- Kluster: 65 , rotationsfaktor: 1.0 , storlek innan zoom: [9, 13]

RA16_27_24.583 i kluster:

- Kluster: 47 , rotationsfaktor: 1.2 , storlek innan zoom: [12, 15]

J11095340-7634255 i kluster:

- Kluster: 47 , rotationsfaktor: 1.6 , storlek innan zoom: [19, 16]

RA16_27_14.489 i kluster:

- Kluster: 47 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 14]

RA16_30_23.390 i kluster:

- Kluster: 47 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [12, 15]

J15530132-2114135 i kluster:

- Kluster: 47 , rotationsfaktor: 0.75 , storlek innan zoom: [13, 16]

RA16_27_38.330 i kluster:

- Kluster: 47 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [12, 13]

GG_Tau i kluster:

- Kluster: 48 , rotationsfaktor: 1.0 , storlek innan zoom: [120, 148]

LupusIII_41 i kluster:

- Kluster: 48 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [27, 37]

G35.03+0.35 i kluster:

- Kluster: 48 , rotationsfaktor: 1.5 , storlek innan zoom: [27, 27]

- Kluster: 64 , rotationsfaktor: 1.6 , storlek innan zoom: [13, 14]

2MASS_J16054540-2023088 i kluster:

- Kluster: 49 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 16]

Haro_6-28 i kluster:

- Kluster: 49 , rotationsfaktor: 0.875 , storlek innan zoom: [16, 15]

2MASS_J16113134-1838259 i kluster:

- Kluster: 49 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [21, 28]

RA16_26_42.854 i kluster:

- Kluster: 49 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [13, 20]

LupusIV_159 i kluster:

- Kluster: 49 , rotationsfaktor: 1.8 , storlek innan zoom: [22, 27]

CX_Tau i kluster:

- Kluster: 49 , rotationsfaktor: 0.875 , storlek innan zoom: [16, 20]

- Kluster: 61 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [26, 23]

J11114632-7620092 i kluster:

- Kluster: 49 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [16, 13]

GM_Aur i kluster:

- Kluster: 50 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [56, 61]

- Kluster: 50 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [44, 49]

- Kluster: 50 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [49, 49]

J05052286+2531312 i kluster:

- Kluster: 50 , rotationsfaktor: 0.875 , storlek innan zoom: [37, 48]

IP_Tau i kluster:

- Kluster: 50 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [28, 22]

RA16_26_40.462 i kluster:

- Kluster: 51 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 19]

VSSG_1 i kluster:

- Kluster: 51 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [26, 23]

D. Bildmanipuleringsinformation

- Kluster: 52 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [27, 27]
HN_Tau i kluster:
- Kluster: 51 , rotationsfaktor: 0.625 , storlek innan zoom: [19, 20]
- Kluster: 51 , rotationsfaktor: 0.625 , storlek innan zoom: [19, 20]
DK_Tau i kluster:
- Kluster: 51 , rotationsfaktor: 0.75 , storlek innan zoom: [22, 16]
RA16_26_23.573 i kluster:
- Kluster: 51 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 19]
V710_Tau i kluster:
- Kluster: 51 , rotationsfaktor: 0.625 , storlek innan zoom: [21, 23]
- Kluster: 51 , rotationsfaktor: 0.625 , storlek innan zoom: [21, 23]
- Kluster: 64 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [16, 27]
LupusIV_150 i kluster:
- Kluster: 51 , rotationsfaktor: 0.625 , storlek innan zoom: [21, 27]
RA16_27_09.086 i kluster:
- Kluster: 51 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]
J16142029-1906481 i kluster:
- Kluster: 51 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 20]
RA16_33_55.596 i kluster:
- Kluster: 51 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 19]
RA16_26_48.468 i kluster:
- Kluster: 51 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]
HD_100453 i kluster:
- Kluster: 52 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [19, 21]
- Kluster: 66 , rotationsfaktor: 1.0 , storlek innan zoom: [70, 61]
- Kluster: 66 , rotationsfaktor: 1.0 , storlek innan zoom: [61, 51]
HD145718 i kluster:
- Kluster: 52 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 15]
J11080297-7738425 i kluster:
- Kluster: 52 , rotationsfaktor: 1.0 , storlek innan zoom: [20, 15]
RA16_27_33.103 i kluster:
- Kluster: 52 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 19]
RA16_27_05.242 i kluster:
- Kluster: 52 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]
V471_Tau i kluster:
- Kluster: 52 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 13]
LupusIII_87 i kluster:
- Kluster: 52 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [21, 22]
J11065906-7718535 i kluster:
- Kluster: 52 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [20, 15]
Sz_54 i kluster:
- Kluster: 52 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [20, 29]
FX_Tau i kluster:
- Kluster: 52 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [23, 19]
RA16_39_45.427 i kluster:
- Kluster: 53 , rotationsfaktor: 1.0 , storlek innan zoom: [19, 16]
RA16_26_48.965 i kluster:
- Kluster: 53 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]
2MASS_J16002612-4153553 i kluster:

- Kluster: 53 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [13, 14]
RA16_25_38.112 i kluster:
- Kluster: 53 , rotationsfaktor: 1.0 , storlek innan zoom: [16, 16]
FM_Tau i kluster:
- Kluster: 54 , rotationsfaktor: 1.6 , storlek innan zoom: [28, 27]
sz84 i kluster:
- Kluster: 54 , rotationsfaktor: 1.5 , storlek innan zoom: [55, 28]
- Kluster: 54 , rotationsfaktor: 1.5 , storlek innan zoom: [43, 23]
LupusIII_103 i kluster:
- Kluster: 54 , rotationsfaktor: 1.0 , storlek innan zoom: [28, 19]
L1489_IRS i kluster:
- Kluster: 55 , rotationsfaktor: 0.6 , storlek innan zoom: [30, 36]
GGD27 i kluster:
- Kluster: 56 , rotationsfaktor: 1.0 , storlek innan zoom: [23, 28]
BHS98_MHO_2 i kluster:
- Kluster: 56 , rotationsfaktor: 1.0 , storlek innan zoom: [40, 37]
BHS98_MHO_1 i kluster:
- Kluster: 56 , rotationsfaktor: 1.0 , storlek innan zoom: [41, 37]
GW_Lup i kluster:
- Kluster: 56 , rotationsfaktor: 0.875 , storlek innan zoom: [35, 34]
2MASS_J04394488+2601527 i kluster:
- Kluster: 56 , rotationsfaktor: 1.0 , storlek innan zoom: [30, 20]
FZ_Tau i kluster:
- Kluster: 56 , rotationsfaktor: 1.0 , storlek innan zoom: [26, 21]
RA16_27_41.736 i kluster:
- Kluster: 57 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [8, 13]
J10563044-7711393 i kluster:
- Kluster: 57 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [27, 20]
J04362151+2351165 i kluster:
- Kluster: 57 , rotationsfaktor: 0.5 , storlek innan zoom: [14, 15]
- Kluster: 57 , rotationsfaktor: 0.5 , storlek innan zoom: [14, 15]
RA16_23_08.791 i kluster:
- Kluster: 57 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [9, 12]
J10581677-7717170 i kluster:
- Kluster: 57 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [26, 21]
2MASS_J11045701-7715569 i kluster:
- Kluster: 57 , rotationsfaktor: 0.5714285714285714 , storlek innan zoom: [12, 12]
J11080148-7742288 i kluster:
- Kluster: 58 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [16, 13]
IRAS_04187+1927 i kluster:
- Kluster: 58 , rotationsfaktor: 0.625 , storlek innan zoom: [16, 19]
- Kluster: 58 , rotationsfaktor: 0.625 , storlek innan zoom: [15, 19]
J11113965-7620152 i kluster:
- Kluster: 58 , rotationsfaktor: 1.2 , storlek innan zoom: [14, 13]
SSTtau_042307.7+280557 i kluster:
- Kluster: 58 , rotationsfaktor: 1.5 , storlek innan zoom: [23, 22]
2MASS_J16270942-2148457 i kluster:
- Kluster: 58 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 12]
2MASS_J16273718-2430350 i kluster:

D. Bildmanipuleringsinformation

- Kluster: 59 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [14, 19]
PBB2002_UScoJ160823.2-193001 i kluster:
- Kluster: 59 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [16, 21]
RA16_26_25.464 i kluster:
- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]
HK_Tau_B i kluster:
- Kluster: 59 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [21, 15]
RA16_28_16.498 i kluster:
- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 19]
RA16_27_39.410 i kluster:
- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]
2MASS_J16072625-2432079 i kluster:
- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 16]
HD139614 i kluster:
- Kluster: 59 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [20, 20]
HD142666 i kluster:
- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 16]
HH_48 i kluster:
- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [28, 15]
LupusIII_1005 i kluster:
- Kluster: 59 , rotationsfaktor: 1.0 , storlek innan zoom: [22, 23]
CM_Cha i kluster:
- Kluster: 59 , rotationsfaktor: 0.875 , storlek innan zoom: [27, 34]
IRAS12500-7658 i kluster:
- Kluster: 59 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [27, 33]
V955_Tau i kluster:
- Kluster: 60 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 12]
WSB_26 i kluster:
- Kluster: 60 , rotationsfaktor: 1.0 , storlek innan zoom: [9, 13]
IRAS_04158+2805 i kluster:
- Kluster: 60 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [15, 13]
RU_Lup i kluster:
- Kluster: 61 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [28, 30]
TMC-1A i kluster:
- Kluster: 61 , rotationsfaktor: 0.6666666666666666 , storlek innan zoom: [30, 36]
L1527_IRS i kluster:
- Kluster: 61 , rotationsfaktor: 0.875 , storlek innan zoom: [55, 27]
Sz_114 i kluster:
- Kluster: 61 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [28, 28]
RA16_23_09.223 i kluster:
- Kluster: 62 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]
RA16_26_42.134 i kluster:
- Kluster: 62 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 15]
RA16_26_37.776 i kluster:
- Kluster: 62 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]
RA16_31_30.866 i kluster:
- Kluster: 62 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]
RA16_27_18.367 i kluster:
- Kluster: 62 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]

RA16_27_16.390 i kluster:

- Kluster: 62 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

RA16_25_24.331 i kluster:

- Kluster: 62 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 15]

RA16_27_09.341 i kluster:

- Kluster: 62 , rotationsfaktor: 1.0 , storlek innan zoom: [13, 16]

LupusI_13 i kluster:

- Kluster: 62 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 16]

RA16_27_41.602 i kluster:

- Kluster: 62 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 15]

Hn_24 i kluster:

- Kluster: 62 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [20, 27]

J16141107-2305362 i kluster:

- Kluster: 63 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 16]

RA16_39_52.910 i kluster:

- Kluster: 63 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 15]

HV_Tau_C i kluster:

- Kluster: 63 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [21, 29]

RA16_26_51.953 i kluster:

- Kluster: 63 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [12, 15]

RA16_23_05.431 i kluster:

- Kluster: 63 , rotationsfaktor: 1.0 , storlek innan zoom: [12, 13]

HD141569 i kluster:

- Kluster: 64 , rotationsfaktor: 1.0 , storlek innan zoom: [9, 9]

ITG_33A i kluster:

- Kluster: 64 , rotationsfaktor: 0.75 , storlek innan zoom: [27, 15]

sigOri_518 i kluster:

- Kluster: 64 , rotationsfaktor: 1.1666666666666667 , storlek innan zoom: [16, 22]

J16001844-2230114 i kluster:

- Kluster: 64 , rotationsfaktor: 0.75 , storlek innan zoom: [9, 14]

2MASS_J04382134+2609137 i kluster:

- Kluster: 64 , rotationsfaktor: 1.0 , storlek innan zoom: [14, 9]

VY_Tau i kluster:

- Kluster: 64 , rotationsfaktor: 0.7142857142857143 , storlek innan zoom: [13, 14]

RA16_27_21.790 i kluster:

- Kluster: 64 , rotationsfaktor: 0.75 , storlek innan zoom: [9, 14]

J11075792-7738449 i kluster:

- Kluster: 64 , rotationsfaktor: 1.3333333333333333 , storlek innan zoom: [14, 9]

2MASS_J10533978-7712338 i kluster:

- Kluster: 64 , rotationsfaktor: 0.5714285714285714 , storlek innan zoom: [12, 13]

Sz_51 i kluster:

- Kluster: 65 , rotationsfaktor: 0.75 , storlek innan zoom: [15, 22]

p104-593 i kluster:

- Kluster: 65 , rotationsfaktor: 1.2 , storlek innan zoom: [9, 12]

J11111083-7641574 i kluster:

- Kluster: 65 , rotationsfaktor: 1.8 , storlek innan zoom: [19, 15]

2MASS_J15354856-2958551 i kluster:

- Kluster: 65 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [9, 12]

J10561638-7630530 i kluster:

D. Bildmanipuleringsinformation

- Kluster: 65 , rotationsfaktor: 0.8333333333333334 , storlek innan zoom: [14, 12]
pds99 i kluster:
- Kluster: 66 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [30, 47]
- Kluster: 66 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [33, 47]
UX_Tau_A i kluster:
- Kluster: 66 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [41, 33]
- Kluster: 67 , rotationsfaktor: 1.1428571428571428 , storlek innan zoom: [43, 36]
LupusIII_1003 i kluster:
- Kluster: 69 , rotationsfaktor: 1.0 , storlek innan zoom: [15, 21]
GI_Tau i kluster:
- Kluster: 69 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [21, 19]
EPIC_203824153 i kluster:
- Kluster: 69 , rotationsfaktor: 1.0 , storlek innan zoom: [36, 23]
2MASS_J16135434-2320342 i kluster:
- Kluster: 69 , rotationsfaktor: 0.8571428571428571 , storlek innan zoom: [14, 15]

INSTITUTIONEN FÖR RYMD-, GEO- OCH MILJÖVETENSKAP
CHALMERS TEKNISKA HÖGSKOLA
Göteborg, Sverige
www.chalmers.se



CHALMERS