

Projektrapport: AI och medborgarforskning för porslinsföremål

Martin Willbo

David Vikstrand

Olof Mogren

Niels Stor Swinkels

My Johansson Dufva

RISE Research Institutes of Sweden
Rörstrand Museum

1 Sammanfattning

I detta projekt har vi arbetat med att förbättra och utöka tidigare maskininlärningslösningar för att modellera porslinsobjekt. Vi har utvecklat modeller som är mer robusta och som kan analysera ytterligare information i bilder av porslinsföremål, till exempel stämplar. Detta har varit möjligt tack vare ett noggrant arbete med att samla in och annotera data.

Vi har samlat in bilder från olika källor, inklusive källor på internet och fotografier tagna av Rörstrand Museum specifikt för projektet. Dessa bilder har använts för att känna igen dekorer, och en del av dem har också annoterats för stämpeligenkänning. Vi har dessutom markerat stämpelpositionen på många bilder för att anpassa en stor, öppet tillgänglig och förtränad modell för objektigenkänning. Modellerna som utvecklades i detta projekt har visat sig fungera tillräckligt bra enligt utvärderingar på nya testdata.

2 Bakgrund

Intresset för Rörstrand är stort både i Sverige och utomlands, och Rörstrand Museum har en viktig uppgift att bevara och sprida kunskap om detta. Tidigare skickades frågor med post tillsammans med framkallade foton, vilket krävde stor ansträngning från den som ville ha svar. Idag skickas frågor via e-post och det är enkelt att ta ett foto, vilket har ökat mängden frågor. Frågorna handlar ofta om signaturer, konstnärer och upplagor. Ibland gäller det föremål som finns i samlingarna och ibland inte. Av de cirka 17 000 föremålen i samlingarna är cirka 4 800 digitaliserade.

Det innebär mycket arbete att hitta svar på frågorna, då det kan innebära att man måste leta i museets magasin och arkiv, samt söka i böcker och digitala källor. Detta arbete tar ungefär 50% av en heltidstjänst. Dessutom är det bara den som ställt frågan som får ta del av svaret. För närvarande besvaras inte inkommande frågor om porslin via e-post, eftersom det saknas resurser för detta.

3 Genomförande

I detta projekt har RISE arbetat med att ta fram en AI-baserad automatiserad assistent för att svara på frågor om porslinsobjekt från Rörstrand. Ett antal AI-modeller har tagits fram och utvärderats och som slutresultat från projektet har ett urval av modeller valts ut och tränats och byggts ihop med ett webbgränssnitt som kan integreras på Rörstrand Museums hemsida.

Vi har fokuserat på tre spår: medborgarforskning, stämpeligenkänning, och dekorigenkänning. Övriga spår från föregående projekt har prioriterats ned, samt ersatts av tabell-uppslagning (dvs, när modellen har kommit fram till en dekor, så slår systemet upp vem som formgivit den samt när, och rapporterar detta).

3.1 AI och datorseende

Modellerna som tagits fram är baserade på djupa neuronnet (deep learning), och mer specifikt faltningssnät (convolutional neural networks; Fukushima (1979)), den AI-teknik som på senare tid visat sig överlägsen inom många tillämpningar inom datorseende. Modellerna tar en bild som indata och kan ge svar om exempelvis designer/konstnär, årtal, dekor och objektstyp. Besluten från modellerna är en avancerad statistisk analys baserat på inlärning från tusentals bilder insamlade från Rörstrand Museums egna digitaliseringsansatser, samt data från andra museum genom samarbetet K-Samsök och Kringla i kombination med andra datakällor (se datainsamling nedan).

Rörstrand Museums egna data hämtades hem från databasen *Carlotta* (museets föremålshanteringssystem), kurerades och sammanställdes till format som är lämpliga för maskininlärning. Datan som hämtats från Carlotta är av god kvalitet, med högkvalitativa bilder och rikligt med information om objekten på bilderna, som konstnär, årtal, dekor, material och objektstyp. Carlottadatan innehåller ungefär 5000 objekt med bilder, samt ca 4000 bilder på stämplor. Data hämtades också från Kringla, där datan är i varierande kvalitet; bilderna har olika bakgrunder och annoteringarna är inte enhetliga. I detta fall har annoteringarna hämtats ut från en textbeskrivning som finns för de flesta objekt. Det ger också en källa till fel, men manuell inspektion visar att datan till stor del är korrekt och att det har hjälpt modellerna att bli bättre på sina analyser. Datan från Kringla innehåller 3500 bilder med annoteringar.

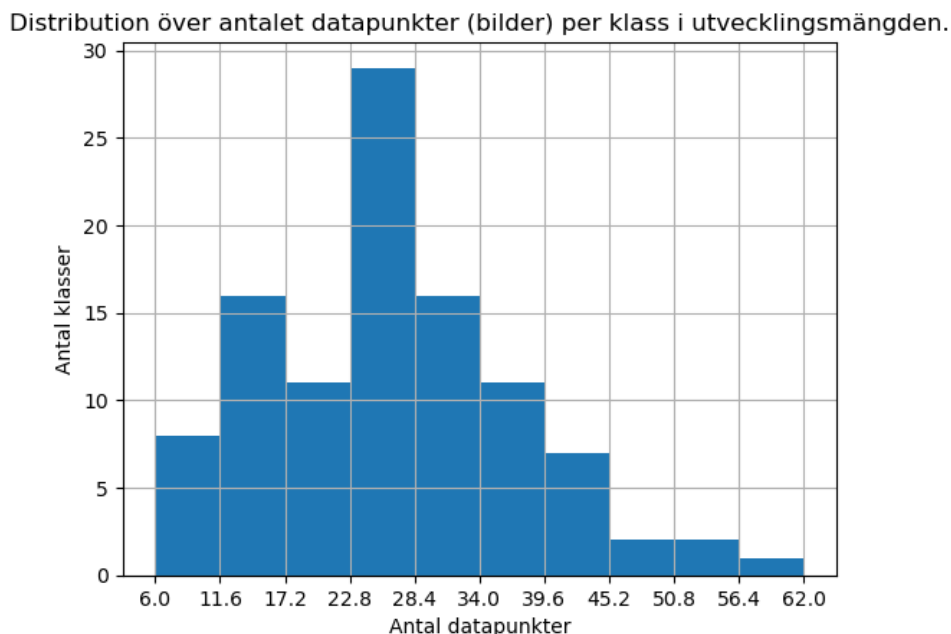
3.2 Begränsningar

För att öka robustheten i modellerna, särskilt för dekorprediktion, valde vi att begränsa antalet dekorer som modellen tränas på och kan känna igen under tester. Genom att minska antalet dekorer som modellen ska kunna förutsäga, blir problemet mindre komplext, vilket gör det möjligt att skapa framgångsrik modellering även när datamängden är begränsad. Rörstrand Museum valde ut 103 dekorer som de ansåg vara viktiga för modellen att kunna särskilja.

3.3 Datainsamling

En slutsats från föregående projekt var att kompletterande data kommer att behövas för att en tillfredsställande kvalitet och robusthet i modellerna ska kunna säkerställas. Utöver de datamängder som samlats ihop under föregående projekt (Carlotta och Kringla, som nämns ovan), så har därför ett antal kompletterande datakällor sammanställts.

Stämpelbilder Vi sorterade ut stämpelbilder från de befintliga datamängderna (Carlotta och Kringla) och annoterade sedan dessa manuellt och halvautomatiskt. Detta för att kunna träna objektigenkänningsmodeller för att detektera stämplor i foton på porslinsobjekt. Totalt annoterades ca 3500 bilder för träning av stämpeldetektormodellen.



Figur 1: Fördelning över antalet datapunkter per dekor, webbilder (notera att denna samling i bilden är kallad utvecklingsmängden).

Bilder från webben Ett relativt stort antal bilder söktes up på webben och användes för att bygga en datasamling med mer variation än de befintliga datamängderna (Carlotta och Kringla). Detta motiverades av att modellen med stor sannolikhet blir mer robust och kommer bättre kunna klassificera användarfotografier då gapet mellan träningsbilder och testbilder blir mindre. Totalt samlades över 2700 bilder in på detta viset med objekt i de 103 valda dekorerna. Se figur 1 för en fördelning över antalet datapunkter per dekor som samlades in från webben (notera att denna samling i bilden är kallad utvecklingsmängden).

Nya fotografier från museet Rörstrand Museum kompletterade de insamlade webbBilderna med egna fotografier över de 103 dekorerna. Detta motiverades av saknaden av ett testdataset som modellerna kan utvärderas på efter träning. Detta är ett moment som är vanligt att utföra inom maskininlärning då det är svårt att säga något om hur modellen kommer att bete sig utan detta momentet. Viktigt är att datamängden som används för detta är så lik den data som modellen ska analysera vid användning, med andra ord användarfotografier på porslinsföremål.

Medborgarforskning Inom medborgarforskningen så har ett nytt webbgränssnitt utvecklats och sju satts. Porslinsfantaster har uppmuntrats att mata in bilder med beskrivningar i ett formulär, och Rörstrand Museum har utropat en tävling med priser till dem som var med och bidrog. Se figur 2. Totalt engagerade detta 45 olika personer som registrerade sig för att bidra med bilder från sina egna samlingar. På detta sätt samlades cirka 100 bilder in.

Övriga källor. Utöver de datamängder så finns det andra museer med Rörstrandsporslin i sina samlingar, till exempel Kulturen i Lund, Röhsska museet och Nationalmuseum. En ansats

för att hämta data även från dessa gjordes, men tyvärr visade det sig inte lönt att göra nytta av datan från dessa museer på grund av följande.

- Informationen i datan var gles.
- Datan var inte enhetlig i sin formatering.

Glesheten i datan betyder att bara viss information fanns angiven om varje objekt. Bristen på enhetlighet i formateringen betyder att datans struktur var begränsad, i många fall behövde information tas ut från fri text vilket är problematiskt, arbetskrävande och leder ofta till fel.

Kulturen i Lund listar 1451 objekt vid en sökning på "Rörstrand". Informationen i databasen har stor variation, många objekt saknar fotografier. Många av objekten har någon information angiven under metadatafälten för "sakord", "location", "time" samt "person". Dessa är dock ifyllda på många olika format och en manuell genomgång hade varit nödvändig för att kunna använda dessa som träningsdata.

Röhsska Museet listar 540 objekt vid en sökning på "Rörstrand". Objekten beskrivs oftast med material, ort och ungefärlig tidspunkt. Det hela är i fritextformat och hade behövt en manuell genomgång för att prepareras för användning som träningsdata.

Nationalmuseum listar 1197 objekt under varumärket Rörstrand. Objekten är från olika tidssåldrar (från 1700-talet till idag), och kan ge en översikt över Rörstrands produktion och utveckling genom tiderna. Detta inkluderar varierade objekt som urnor och vaser och mer vardagliga föremål som tallrikar och skålar. Objekten i samlingen har ofta metadata som inkluderar tillverkningsår, artistinformation, material och tekniker. Dock hade det behövt göras en manuell genomgång av varje objekt, då det inte är standardiserat hur denna data är formaterat. Många fält innehåller ofta flera olika metadata separerade med semikolon.

Ett beslut togs motiverat på tekniska grunder om att istället lägga större fokus på att arbeta med datakällor som enklare kunde ge större bidrag till bättre resultat, vilket resulterade i spåren med bilder från webben och nya fotografier från museet, se ovan.

3.4 Stämpeligenkänning

För stämpeligenkänningsmodellen valdes som bas en förtränad dino-modell. Det är en stor modell som har tränats på stora mängder data. Detta är ett vanligt förfarande inom maskininlärning, att man tar brett välfungerande modeller och anpassar dem på mindre, mer domän-specifika, datamängder. I detta fallet anpassades modellen till uppgiften stämpeligenkänning med den manuellt annoterade stämpeldatamängden.

Att känna igen var en stämpel finns i en bild av ett porslinsobjekt var en del av det som skulle modelleras i projektet. En annan önskad funktion för stämpeligenkänningsmodellen var att kunna förutsäga vilket årtionde objektet kommer ifrån baserat på en Rörstrand-stämpel. För denna uppgift användes en mindre modell. Denna mindre modell analyserar de områden som den större stämpeligenkänningsmodellen hittar. Modellen är förtränad och anpassas med hjälp av den manuellt annoterade stämpeldatamängden i kombination med data om objektens årtionden.

LADDA UPP

Hjälp oss att träna AI genom att ladda upp bilder på ditt eller dina Rörstrandsföremål. Fyll i så mycket information om föremålet som du kan, men enbart information som du är säker på. De fält du inte vet lämnar du tomt.

OBS: Fyll i endast ett föremål i taget. Separera även kopp och fat som olika föremål.

The screenshot shows a web form for uploading items. It includes several sections with instructions:

- FÖREMÅLSBILDER**: A button to 'Ta / välj kort' and an instruction box stating: 'Lägg till en eller flera bilder av ditt föremål. Ta bilden så att föremålet är skarpt och välbelyst, men använd ej blixt. Försök att få med hela föremålet i bilden samtidigt som detaljer som till exempel dekormönster eller andra utsmykningar som guldkanter är tydliga.'
- STÄMPELBILDER**: A button to 'Ta / välj kort' and an instruction box stating: 'Lägg till en eller flera bilder på stämpeln på undersidan av ditt föremål. Ta bilden så att stämpeln på undersidan är skarp och tydlig.'
- TYP**: A dropdown menu with 'Okänd' selected and an instruction box: 'Vad är det för typ av föremål? Välj en typ ur listan. Du kan söka i listan genom att börja skriva i rutan. Om rätt typ inte finns i listan, skriv i rutan och tryck enter.'
- DEKOR**: A dropdown menu with 'Okänd' selected and an instruction box: 'Vad heter dekoren på föremålet? Välj en ur listan. Du kan söka i listan genom att börja skriva i rutan. Om rätt dekor inte finns i listan, skriv i rutan och tryck enter.'
- DESIGNER(S)**: Two dropdown menus with 'Okänd' and 'sOm' selected, and an instruction box: 'Vem har designat föremålet? Lägg till en eller flera designers, och om möjligt vad de hade för roll i designen. Du kan söka i listan genom att börja skriva i rutan. Om rätt designer inte finns i listan, skriv i rutan och tryck enter.'
- ÅRTAL**: A text input field with 'Årtal' and a 'Fyll i som period' button, and an instruction box: 'När var föremålet i produktion? Ange ett årtal. Du kan också välja att fylla i start- och slutår för att ange en tidperiod.'

At the bottom of the form is a blue button labeled 'Ladda upp'.

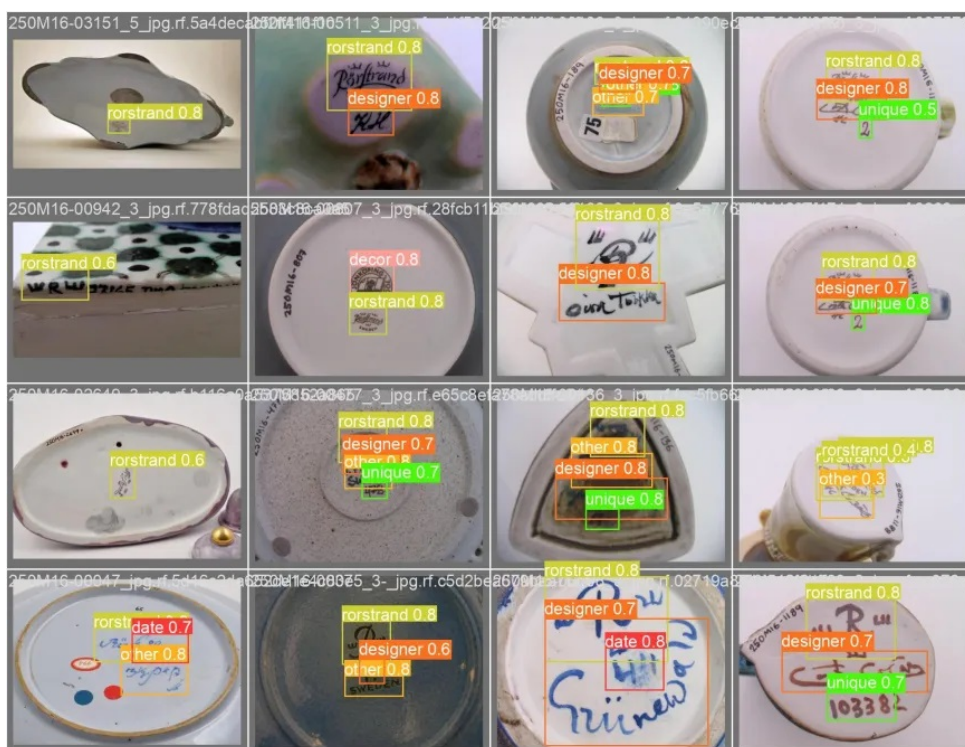
Figur 2: Gränssnittet som utvecklats för insamling av frivilligas bilder.

3.5 Dekorigenkänning

För dekormodelleringen valdes som bas en förtränad bildklassificerare. Den är tränad på ett brett sortiment av fotografier. Denna modell har sedan anpassats för uppgiften dekorigenkänning med hjälp av webbBilderna.

För att möjliggöra en potentiell målrumsexpansion (utöka antalet dekorer som modellen kan känna igen efter parameteruppdatering) vid ett senare skede togs beslutet att sedan använda denna anpassade dekormodell i en prototypical network (?) uppsättning. Det vill säga, dekormodellen klassificerar fotografier genom att mäta avståndet mellan abstrakta representationer för dem istället för att explicit separera via linjära transformationer.

Modellen valdes sedan i slutändan med hjälp av en valideringsmängd. Denna sammansattes genom att slumpmässigt välja datapunkter från de 103 dekorer på webbBildsmängden. Återstående datapunkter utgjorde träningsmängden, som var ca 80% av alla webbBilderna. Efter modellvalet utvärderades den också på en testmängd, som i detta projektet är de nya fotografierna från muséet.



Figur 3: Exempel av de prediktioner som kommer ur steg ett av stämpeligenkänningsmodellen. Här ser man vart modellen tror sig att det finns till exempel Rörstrandsloggor i bilderna.

3.6 Ensemble

Man använder ensemblemodeller inom maskininlärning för att förbättra prestandan och robustheten hos prediktionsmodeller. Ensemblemetoder kombinerar flera individuella modeller för att skapa en sammansatt modell som ofta presterar bättre än någon enskild modell. Man får detta oftast genom att introducera skillnader under träning för dessa individuella modeller, till exempel genom att visa dem slumpmässigt valda delmängder av träningsmängden eller att introducera data i en annan ordning eller att modeller börjar från olika ställen i parameterrummet.

I detta projekt använder vi en ensemblemodell främst för de fördelar man får i robusthet. När flera modeller gemensamt måste rösta fram vilken prediktion som ska presenteras har det en tendens att minska antalet felprediktioner man gör totalt, främst för att på den datapunkten felande modeller blir överröstade av sina korrekta motsvarigheter.

3.7 Webgränssnitt

Slutresultaten från detta projektet är huvudsakligen modeller som är tränade på de nyinsamlade datamängderna. Därmed går det att återanvända mycket av det gränssnitt som var ett resultat av det tidigare projektet. Ett arbete har gjorts för att sammanfoga det tidigare webgränssnittet med medborgarforskningsplattformen för att det helha ska vara tillgänglig från en plats. Dessutom har en viss finslipning av resultatpresentationen gjorts samt en del kodarbete på baksidan utav gränssnittet för att möjliggöra tillägg av stämpeligenkänning och prediktion.

4 Projektresultat

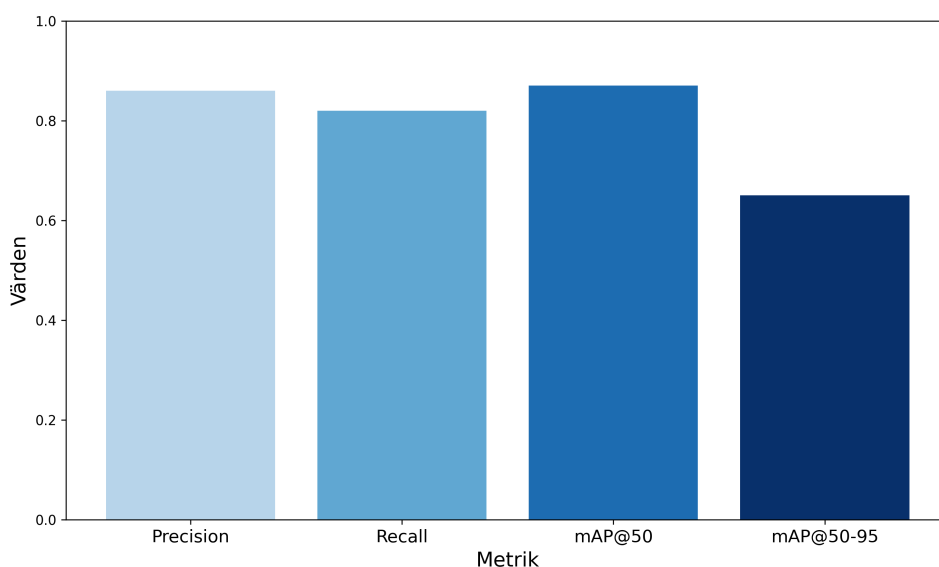
Här presenterar vi resultaten för de olika modelleringsuppgifterna i projektet. Först visas hur väl stämpeligenkänningsmodellen hittar stämplat och sedan hur den klassificerar dessa om de hittas. Sist visas träffsäkerheten för dekorigenkänningsmodellen och hur spridning av resultaten har skett.

4.1 Stämpeligenkänning

Projektresultatet för stämpeligenkänningsmodellen består av två komponenter då modellen är tvådelad. En del som predicerar vart någonstans i bilden som stämpeln befinner sig (tex Rörstrandslogga) och en komponent som sedan predicerar årtionde från det område i bilden som är predicerat av det första steget.

Resultatet för det första steget, alltså att predicera var i bilden en stämpel befinner sig visas i figur 3. Vi har fyra mått: precision, recall, mAP@50 och mAP@50-95. Precision berättar att **86%** av de objekt som identifieras är korrekta. Recall betyder att modellen framgångsrikt hittar **82%** av alla faktiska objekt i en bild. mAP@50 mäter modellens noggrannhet med en standard-tröskel, och visar att noggrannheten i vår modell är **87%**. mAP@50-95 beräknar noggrannheten över flera trösklar, vilket kan ses som ett striktare mått.

Sammantaget tyder dessa siffror på att vår modell är bra på att upptäcka stämplat med hög noggrannhet och tillförlitlighet. Designersignaturer är en svårare utmaning, vilket drar ner de genomsnittliga måtten något. Allt som allt visar detta att vår modell är ett starkt verktyg för delproblemet att känna igen och identifiera stämplat på Rörstrandsobjekt.



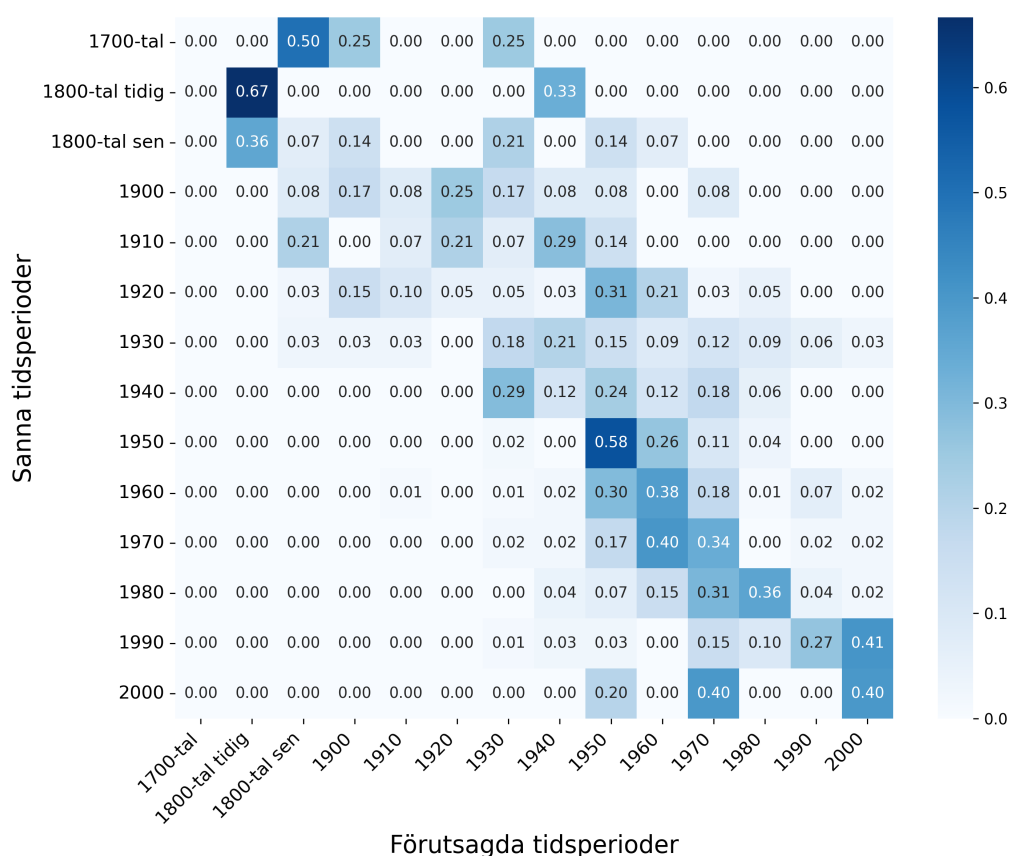
Figur 4: Resultat för steg ett av stämpeligenkänningsmodell.

I figur 5 visar vi resultatet för det andra steget, där vår modell avgör vilken tidsperiod ett objekt är ifrån. Detta är baserat på ett urklipp av en bild där steg 1 beslutat att det finns en Rörstrand-logotyp. Figuren visar en förväxlingsmatris som illustrerar hur bra vår klassificeringsmodell

presterar. Matrisen jämför de sanna tidsperioderna för objekten med de förutsägelser som modellen har gjort. Diagonalcellerna i matrisen visar antalet korrekta förutsägelser för varje tidsperiod, medan de icke-diagonala cellerna visar antalet felaktiga förutsägelser.

För vår modell vill vi se höga värden i celler som är så nära diagonalen som möjligt, detta betyder att modellen korrekt förutsatt den sanna tidsperioden för objektet. Observera att den tidsperiod ett objekt tillhör kan sträcka sig länge än dess sanna tidsperiod, vilket betyder att en utmatning till höger om diagonalen fortfarande kan vara korrekt.

Sammanfattningsvis tyder dessa siffror på att vår modell presterar bra och gör korrekta förutsägelser för det mesta. De höga värdena i diagonal- och närliggande celler indikerar en överensstämmelse mellan de sanna etiketterna och modellens förutsägelser. Detta gör vår modell till ett givande redskap för att identifiera och klassificera tidsperioder.



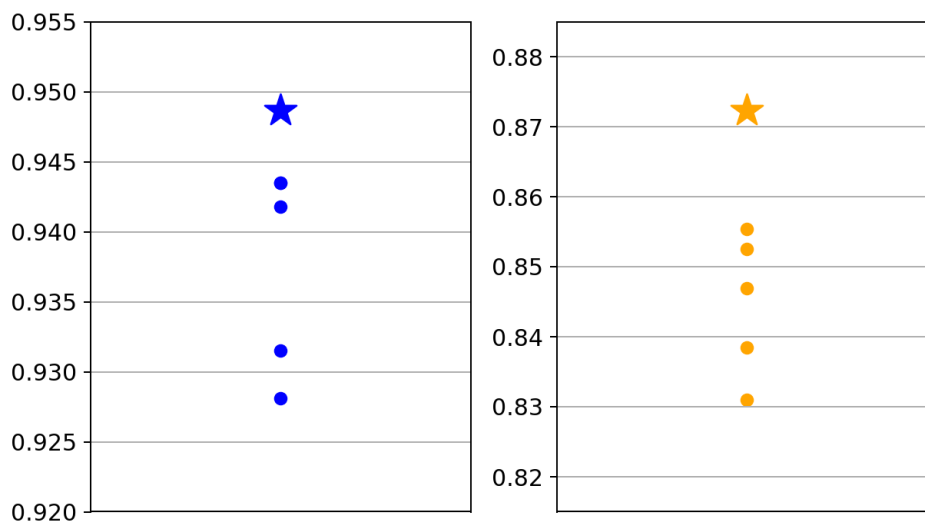
Figur 5: Resultat för steg två av stämpeligenkänningsmodell.

4.2 Dekorigenkänning

I denna sektion presenteras resultat på dekorigenkänning för valideringsmängden som modellen valdes på och testmängden bestående av de fotografier Rörstrand Museum har tagit. I figur 6 ser vi dekormodellens träffsäkerhet, som är andelen rätt över hela mängden, på både validerings- och testmängden. Figuren visar träffsäkerheten för enskilda modeller och för den ensemble som är en sammanslagning av dem. Det är tydligt att en ensemble som består av

flera enskilda modeller presterar bättre än de enskilda modellerna för sig, för respektive mängd. Ensemblemodellen når ca 95% i träffsäkerhet på valideringsmängden, detta är ett mycket starkt resultat.

Dock, av större intresse är ensemblemodellens uppmätta träffsäkerhet på testmängden, ca 87%. Testmängden skiljer sig åt från de insamlade webbBilderna, i det att de än mer efterliknar de porslinsfrågor som Rörstrand Museum mottagit, och är därmed också ett prov i modellens förmåga att vara robust mot en bredare kategori av data.



Figur 6: Resultat för dekorigenkänningsmodell. Vänster i blått: Valideringsmängd. Höger i oranget: Testmängd (Museets nya fotografier). Enskilda modeller och deras träffsäkerhet på respektive mängd efter träning representeras med punkter. Ensemblemodellen, som är en sammansättning av dessa enskilda modeller, representeras med en stjärna. Notera att skalorna i de två figurerna (vänster, höger) inte är desamma.

4.3 Spridning av resultat

Som avslutande moment på projektet hölls ett slutseminarie i rörstrandmuséums lokaler i Lidköping där 13 personer deltog på plats och 28 personer som deltog digitalt. På slutseminariet presenterades webgränssnittet och hur man interagerar med det för att antingen bidra till medborgarforskningsmängden eller för att få prediktioner om sina porslinsföremål.

Efter presentationen fanns det ett antal porslinsföremål tillgängliga för deltagare att använda som exempel när de utforskade webgränssnittet och dess funktioner.

Utöver slutseminariet har projektet även genererat inbjudningar att presentera resultaten vid andra tillfällen. Under hösten 2024 blev My Johansson Dufvainbjuden att tala på en av regionens museinätverksträffar samt delta i ett panelsamtal på Digikult i början av oktober.

En manual har skrivits med en beskrivning av hur man kan återskapa resultaten i detta projekt, eller hur man kan skapa en liknande modell för andra samlingar.

Den insamlade datan och den källkod som använts för träning och användning av modeller

finns tillgänglig.

5 Diskussion och slutsatser

Detta projekt har haft som mål att stärka den maskininlärningslösning som ligger till grund för det webgränssnitt för porslinsfrågor som Rörstrand Museum vill tillhandahålla. Lösningens robusthet mot skillnader i indata behövde stärkas, då det i tidigare projekt visade sig vara stor skillnad på data som användes för träning och den typiska användardatan. Utöver detta behövde även porslinsföremålets stämplat undersökas som potentiellt beslutsunderlag.

Just dataskillnaden i det tidigare projektet har till stor del varit ledande för projektutförandet. Det bestämdes tidigt att data av en annan karaktär än den som användes för träning i den tidigare iterationen behövdes. Detta ledde till en bildinsamling och fotografering som resulterat i sorterade mängder med data som kan användas vid fortsatt arbete med porslinsmodellering. Detta gäller även den datamängd över annoterade bilder på undersidan av porslinsföremål.

Under projektets gång har det utvärderats flera tillvägagångssätt med målet att utveckla effektiva och robusta maskininlärningslösningar. Det arbetet har gett resultat och lett till den slutgiltiga produkten som ger bättre resultat än den tidigare. Arbetet med datainsamling och annotering gett stora förbättringar, och arbetet med modellering har bidragit med ytterligare små förbättringar av robustheten.

Den modell som togs fram i projektet är begränsad till 103 olika dekorer, och kan i många fall ge information om tidsperiod och konstnär, även för objekt som har andra dekorer. Rörstrand Museum ser detta som en prototyp och en teknikdemonstration men en som hittills inte är tillräckligt mogen för att tillhandahålla som en tjänst från museets hemsida. Begränsningarna togs fram för att kunna skapa robusta modeller utifrån den datatillgång som fanns tidigare tillsammans med vad som kunde tas fram i detta projekt. En framtida lösning skulle kunna tas fram som är tränad på större mängder data och som därmed har mindre begränsningar.

Generellt kan sägas om projektresultatet att vi visat på många svårigheter som finns i den här typen av problemuppställning. Resultaten har inte mött förhoppningarna för att kunna skapa en artificiell assistent med den kompetens som skulle ha behövts. En högre träffsäkerhet hade önskats på föremål som finns representerade i träningsdatan, och man hade även önskat en större generalisering, med en nyanserad kunskap som idag inte kan uppnås med maskininläring i kombination med den data som tagits fram och använts i detta projekt.

En manual har skrivits med en beskrivning av hur man kan återskapa resultaten i detta projekt, eller hur man kan skapa en liknande modell för andra samlingar. Det är dock svårt att säga hur mycket arbete detta kommer att kräva då det beror på hur den tillgängliga datan ser ut.

6 Framtida arbete

Detta projektets slutresultat pekar mot att det genomförande som presenterats i denna rapport kan användas som underlag för framtida arbete kring modellering av porslinsobjekt. Framförallt har datainsamlingen som har genomförts för flera olika typer av relaterad bilddata under projektets gång har visat på vikten av gedigna datamängder för ökad robusthet i test.

Dekorigenkänningsmodellen kan med ett mindre arbete utökas så att den kan särskilja ett större antal dekorer, även efter träning av representationsdelen av modellen. Detta är något som kan vara en framtida utökning om nya datamängder samlas in för ytterligare dekorer utöver de 103 som har arbetats med i detta projekt.