Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Design and development of multilayer cotton masks via machine learning

Y. Leow a, c, J.K. Shi b, c, W. Liu b, X.P. Ni a, P.Y.M. Yew a, S. Liu a, Z. Li a, Y. Xue b, *, D. Kai a, **, X.J. Loh a, * a Institute of Materials Research and Engineering (IMRE), Agency for Science, Technology and Research (A*STAR), 2 Fusionopolis Way, Innovis, #08-03, Singapore, 138634, Singapore
b Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, Connexis South Tower, #21-01, Singapore, 138632, Singapore
c Both authors contributed equally to the work.

ABSTRACT

With the ongoing COVID-19 pandemic, reusable high-performance cloth masks are recommended for the public to minimize virus spread and alleviate the demand for disposable surgical masks. However, the approach to design a high-performance cotton mask is still unclear. In this study, we aimed to find out the relationship between fabric properties and mask performance via experimental design and machine learning. Our work is the first reported work of employing machine learning to develop protective face masks. Here, we analyzed the characteristics of Egyptian cotton (EC) fabrics with different thread counts and measured the efficacy of triple-layered masks with different layer combinations and stacking orders. The filtration efficiencies of the triple-layered masks were related to the cotton properties and the layer combination. Stacking EC fabrics in the order of thread count 100-300-100 provides the best particle filtration efficiency (45.4%) and bacterial filtration efficiency (98.1%). Furthermore, these key performance metrics were correctly predicted using machine-learning models based on the physical characteristics of the constituent EC layers using Lasso and XGBoost machine-learning models. Our work showed that the machine learning-based prediction approach can be generalized to other material design problems to improve the efficiency of product development.

© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The outbreak of the novel SARS-CoV-2 virus has brought about a global COVID-19 pandemic, threatening humanity and putting the world at a standstill. Given the interconnectedness of the current society, rapid human-to-human transmission of the virus can occur through respiratory droplets or physical contact [1–8]. With rapid viral mutation and limited vaccine coverage, preventive measures are indispensable in preventing disease escalation. To curb the spread of the SARS-CoV-2 virus, many countries have imposed widespread lockdowns and travel restrictions, in addition to social distancing and contact tracing measures [4,9]. As a consequence, economies around the world have been adversely affected.

Among various safety measures, masks are very effective in reducing the spread of SARS-CoV-2 with a very low implementation cost. To mitigate the demand for surgical masks, simple masks are easily made with common household materials as an alternative. Studies have investigated the effectiveness of homemade masks made from cotton, linen, silk, etc., confirming their protection against COVID-19 [10]. Konda et al. evaluated multilayered cloth masks and demonstrated that hybrid face masks made from cotton-silk, cotton-chiffon, and cotton-flannel exhibited filtration efficiencies greater than 80% for particles less than 300 nm and more than 90% for particles more than 300 nm [11], thus highlighting the feasibility of adopting common fabrics as protection against aerosol particles transmission.

However, the development of an efficient face mask requires repeated experiments, making it time-consuming and costly. The materials are also selected based on human intuition or experience.

https://doi.org/10.1016/j.mtadv.2021.100178
2590-0498/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
In this study, we use machine learning to design multilayer masks to significantly reduce the number of experimental measurements. The capabilities of machine learning have been well established for material development in other areas. Thus, it could potentially be an effective tool to guide the design and advancement of masks. Tan et al. successfully optimized the construction of a polyetherimide ultrafiltration membrane through hybrid models built on a backpropagation neural network and genetic algorithm [12]. Predictions obtained from the hybrid models were adequately precise and the most desired membrane exhibits performance within the forecast. Another study by Chi et al. evaluated the use of four different supervised learning approaches for the modeling of poly(vinyl chloride)/polyvinyl butyral ultrafiltration membranes [13]. They concluded that the neural network model is preferable which was further confirmed with experimental data.

In this work, we employed simulated computing studies via artificial intelligence to provide insight into the relative importance of different experimental observables in determining the desired mask performance. In particular, Egyptian cotton (EC), a household fabric material, was used to make multilayer masks with the relationship between fabric properties and mask performance investigated. We hypothesize that a multilayer cotton mask with the targeted filtration efficiency can be designed by knowing the material properties of single-layer fabrics. The physical characteristics (morphology, surface charge, surface area, and pore size) of EC fabric with different thread counts were examined. Triple-layer masks were prepared by stacking the EC fabrics in varying combination of threes can be specified to a hyphen in between the representative number. For example, fabric stacking combination representation number. For example, fabric stacking combination 3-2-1 is interpreted as a fabric layering of EC300 (outer layer), EC200 (middle layer), and EC100 (inner layer).

2.2. Physical characterization of EC fabric

EC fabric specimens were divided into 12 cm by 12 cm squares, and the fabric thicknesses were measured using a Vernier caliper. The specimens were also examined morphologically using a scanning electron microscope (SEM) (JSM6700F, JEOL, Japan). To obtain fiber diameters of different EC samples, a small area of each specimen was cut and sputtered with thin layers of gold before SEM imaging.

To identify the mean flow pore size, pore size distribution, and maximum pore size of EC fabric specimen, pore size-related measurements were conducted based on gas—liquid porometry using a capillary flow porometer (Microflow Porometer, Porous Materials Inc., USA). The method of operation applied was the ‘wet up/dry down’ process. Each specimen was prepared into 3 cm by 3 cm squares and initially wetted with a low surface tension (15.9 dyn/cm) Galwic solution for all voids to be filled. Afterward, it was fitted into the sample chamber securely with an O-ring. A gradual vertical air pressure was introduced to push the Galwic solution out. Pressure and flow throughout the experimental procedure were measured and graphed. The maximum pore size can be calculated based on the bubble point, which is identified as the pressure when the Galwic solution first escapes the pores. The mean pore size and pore size distribution were also derived from the porometer. The resultant pore size quantification obtained was on the micron scale.

An additional examination was carried out to determine the total specific surface areas and fiber pore size. The characteristics of individual EC fabrics were determined by applying the Brunnauer—Emmett—Teller (BET) theory with nitrogen gas as the adsorbate. The specimens are subjected to increasing partial pressure of nitrogen gas and the specific surface area is calculated when a monolayer of gas covering all surfaces is formed on the substrate at a definite pressure. As the gas pressure increases past the definite pressure, calculation of average pore size is possible on the assumption that all pores are permeated with nitrogen gas molecules. Before measurement with an adsorption analyzer (ASAP 2020 Plus, Micromeritics, USA), each specimen of approximately less than 0.9 g was oven-dried to remove all water content.

The surface charge of each EC fabric variation was quantified by performing surface zeta potential analysis with SurPASS 3 instrument from Anton Paar. Individual samples were mounted onto a cylindrical cell specified for fibers, granular and powder samples. For each measurement, pH titration started at a native pH of 0.001 M potassium chloride (KCl) electrolyte solution and was adjusted using 0.05 M hydrochloric acid (HCl). Zeta potential as a function of pH was investigated while varying the pH of KCl solution from 2 to 6.

2.3. Breathability and filtration efficiency measurement

Breathability assessment was conducted by differential pressure testing. Cross-sectional area specimens measuring 2.5 cm in diameter were prepared for the different EC variants. All specimens were conditioned at (21 ± 5) °C and (85 ± 5)% relative humidity until atmospheric equilibrium. Sample testing and experimental setup (Fig. S1) were conducted in accordance with EN 14683 standards with airflow adjusted to 8 L/min. Differential pressure, ΔP, for each EC variant was attained by taking the pressure difference of individually recorded pressure before and after air flows through the cross-sectional area of the specimen.

The airborne pollutant filtration capability of the EC fabrics was studied using the particle filtration efficiency (PFE) test. The experimental setup (Fig. S2) and tests were conducted according to the ASTM F2299 standard. All samples were tested under fixed conditions with a monodisperse particle size of 0.1 μm at 10 cm/s airflow for 1 min of sampling time. Five upstream and downstream aerosol counts were obtained and averaged per EC sample. PFE is calculated using the following formula (Eq. (1)).

\[
PFE = \left(1 - \frac{M_d}{M_u}\right) \times 100\%
\]

where \(M_d\) is the downstream particle counts concentration and \(M_u\) is the upstream particle counts concentration.

Evaluation of biological filtration capability for EC fabrics was conducted using bacterial filtration efficiency (BFE) tests.
conforming to the ASTM F2101–19 standard, at a flow rate of 28.3 L/min and across a cross-sectional area of 50 cm². Gram-positive bacterium Staphylococcus aureus was selected in this study for its clinical relevance and cause of nosocomial infection in healthcare facilities [14]. The mean particle size of the spherically shaped bacteria is about 3 ± 0.3 μm. A control was also prepared under identical conditions without EC fabric filter media. BFE is calculated using the following equation (Eq. (2)).

\[ \text{BFE} = \frac{C - F}{C} \times 100\% \]

where C is the average plate count total for test controls and F is the plate count total for a test sample.

2.4. Machine learning

To predict the three key performance metrics ΔP, PFE, and BFE, of multilayer cotton masks, we can construct the features from either single-layer ΔP, PFE, and BFE, or single-layer material characteristics as shown in Table 1. The latter approach was chosen because it requires less experimental effort to generate the training data, thus being more time-cost efficient. Focusing on more fundamental properties also makes this method more generalizable to new cotton materials.

There are seven physical characteristics of single-layer cotton that we consider are the most important in determining ΔP, PFE, and BFE:

- Fabric thickness
- Fiber diameter
- Fiber strand diameter
- Zeta potential
- Specific surface area
- Surface pore size
- Mean pore size

With the additional degree of freedom from the layer index (3 for a three-layer mask), there are 7 × 3 = 21 features in total.

Given the small data size, we compared two regularized machine-learning models: Lasso (Least Absolute Shrinkage and Selection Operator) and XGBoost (Extreme Gradient Boosting).

Lasso is an L1-norm regularized least-squares linear regression algorithm with sparse solutions [15,16]. The objective function (Eq. (3)) to be minimized is

\[ \min_w \frac{1}{2n_{\text{samples}}} ||Xw - y||_2^2 + \alpha ||w||_1, \]

with X being the features, y the target, α a constant, and w the L1-norm of the coefficient vector. The first term is the square loss function that measures the difference between the predicted and the actual target value, while the second term is the regularization term that penalizes coefficient w with large absolute values. With L1-norm regularization, coefficients are often pushed to exactly zero during gradient descent in training, leaving only a few non-zero coefficients, effectively acting as a feature selector. The features are ranked in descending orders of the magnitude of w to get the feature importance.

On the other hand, XGBoost is a boosted decision tree-based machine-learning model [17]. It uses gradient boosting on an ensemble of decision trees and combines a number of weak learners to create a strong model [18]. One of the advantages is its robust performance in giving accurate predictions on a small training data set. Decision trees are graphs in tree form, with each node representing a condition that splits data into different predicted target values. The optimal splitting conditions are determined by minimizing the entropy after each split or maximizing an information gain. The extent to which a feature contributes to each information gain can be used to determine the feature importance. Effectively, the ‘gain’ measures the relative contribution to the corresponding feature to generating a prediction [19]. It is thus most appropriate to calculate the feature importance based on gain.

Standard k-fold cross-validation was used for hyperparameter tuning in either model.

3. Results

3.1. Characteristics of a single-layer EC fabric

Details on physical characteristics and measurements of single-layer EC fabrics are presented in Table 1. The thickest EC fabric belongs to EC100, 0.182 mm, followed by EC200, 0.128 mm, and lastly EC300, 0.120 mm. The fabric thicknesses for EC200 and EC300 are fairly similar. Surface morphologies of the different EC specimens at ×100 and ×500 magnifications are examined by SEM and represented in Fig. 1. The threads were weaved in a systematic order and each fiber thread comprises multiple thin fibers bundled together as observed in Fig. 1. The width of each bundle, also known as fiber strand diameter, varies across different EC samples. An individual thin fiber has diameters of 12.1, 10.8, and 9.8 μm for EC100, EC200, and EC300 respectively. In the same specimen order, fiber strand diameters are 146, 137, and 103 μm. Since every fiber strand is made up of thin fibers bound together, the decreasing trend for fiber diameter and fiber strand diameter is consistent. The EC fabrics were categorized according to their thread count, which is directly influenced by the fiber diameter. A higher thread count rating means that the fabric is produced with threads of a smaller fiber diameter. This was evident as both fiber and thread became thinner with increasing thread count.

Quantitative information on pore size measurements was obtained from the porometer and tabulated in Table 1 and Table S2. The pore size distribution is also represented in Fig. S5. These results provided insights into the maximum and mean pore size in each distinct EC fabric. Results showed maximum pore sizes of 64.9, 79.2, and 43.5 μm for EC100, EC200, and EC300 specimens respectively. A mean pore size of 6.8 μm for EC100, 14.9 μm for EC200, and 11.5 μm for EC300 were also identified.

According to the BET results in Table 1, the corresponding specific surface areas of EC100, EC200, and EC300 fabrics are 0.59, 1.03,
and 0.71 m²/g. An increasing trend was observed for surface pore size, starting with EC100 at 11.23 nm, then EC200 at 34.94 nm, and lastly EC300 at 53.61 nm.

Surface chemistry analysis was carried out via Fourier transform infrared (FT-IR) spectroscopy and X-ray photoelectron spectroscopy (XPS). The FT-IR and XPS results for EC100 are detailed in Fig. 1(d) and (e) respectively. The remaining FT-IR and XPS results are depicted in Fig. S3, Fig. S4, and Table S1. From the FT-IR spectra, cellulose can be determined to be present in the composition. By comparing the surface chemistry of the three fabrics, there is no significant difference between them. This indicates that the fabrics were of the same material and have largely similar surface chemistry. Hence, the effect of surface chemistry was not considered in this study.

The zeta potentials of the three different fabrics are included in Table 1. The surfaces for all the fabrics are negatively charged. The zeta potentials of EC100 and EC300 are almost the same between –10.2 and –9.8 mV. In particular, EC200 exhibits a zeta potential doubled of both EC100 and EC300 at –20.7 mV.

The essential parameters of single-layer EC fabric masks, breathability and filtration efficiencies, were measured. The outcomes are reported in Table 2. ΔP for EC100 is significantly higher at 118.3 Pa/cm² compared to EC200 and EC300 at 26.6 and 44.8 Pa/cm². Individually, EC100, EC200, and EC300 fabrics had PFEs of 22.8, 9.9, and 17.6%, respectively. PFEs of EC100 and EC300 were almost doubled compared to that of EC200. In the same sequence, the fabrics have BFEs of 84.17, 50.60, and 48.27%, respectively. EC100 has the highest BFE, whereas EC200 and EC300 both have similar BFE results.

3.2. Filtration efficacy of a triple-layer EC fabric

In Singapore’s context, the Health Science Authority (HSA) has recommended surgical masks to have a PFE of at least 80% and a BFE of above 95%. Compared to conventional surgical-grade face masks, PFE and BFE of the single-layer EC fabrics vary in a wide range and were significantly lower than the standards. Hence to improve filtration efficiency and make suitable EC face masks, stacking of different EC fabrics in varying combinations of threes is explored.

Table 3 presents the different combinations of a triple-layered
EC fabric with their corresponding experimental $\Delta P$, PFE, and BFE values. It is evident from Table 3 that both PFE and BFE significantly improved when the EC fabrics are stacked together in threes. Consequently, the stacking combination of EC fabrics also led to an increase in $\Delta P$, up to between 67.5 and 253 Pa/cm². PFE values ranged between 28.6 and 46.8%, which falls short of the recommended PFE by HSA. Nevertheless, the improvement was still noteworthy. Excluding the 1–1–1 combination, the best PFE result was achieved by stacking a combination of 1–3–1 at 45.4%. On the other hand, the BFE of the stacked EC fabrics varies from 89.7 to 98.1%. The most notable BFE result, 98.1%, again belongs to the 1–3–1 stacking combination. Six of the tested sample combinations were able to attain over 95% BFE, hence suggesting the potential and suitability of these sequences for homemade EC fabric masks. The increase in all three face mask properties can be attributed to thickness as most of the predictions are still considered within the acceptable range given intrinsic experimental uncertainty and variability.

Table 3
$\Delta P$, PFE, and BFE of triple-layer EC specimens (e.g. 3–2–1) representing the stacking order of EC300, then EC200, and then EC100) obtained experimentally. For the combination 1–2–3, internal means that the flow direction is 1 to 3, while external means that the flow direction is 3 to 1. Sym—triple-layer stacking order is symmetrical in both directions (e.g. 1–2–1), hence there is no BFE differences between the directions of bacterial flow.

| Sample combination | $\Delta P$ (Pa/cm²) | PFE (%) | BFE internal (%) | BFE external (%) |
|-------------------|---------------------|---------|-----------------|------------------|
| 1–1–1             | >253                | 46.8 ± 0.6 | 95.5 ± 0.5 | sym              |
| 1–2–1             | 228.6 ± 7.7         | 43.2 ± 1.1 | 96.7 ± 1.5 | sym              |
| 1–3–1             | 232.2 ± 4.1         | 45.4 ± 1.4 | 98.1 ± 0.8 | sym              |
| 1–3–3             | 138.8 ± 3.4         | 41.9 ± 1.6 | 95.5 ± 0.4 | 92.5 ± 0.1      |
| 2–1–2             | 142.0 ± 7.5         | 38.0 ± 0.4 | 92.3 ± 0.7 | sym              |
| 2–2–3             | 67.5 ± 1.6          | 30.2 ± 0.1 | 93.9 ± 0.3 | 92.7 ± 0.9      |
| 3–1–3             | 142.1 ± 3.4         | 42.5 ± 2.2 | 91.7 ± 0.4 | sym              |
| 3–2–1             | 154.1 ± 5.3         | 38.5 ± 2.1 | 94.3 ± 0.2 | 95.6 ± 0.3      |
| 3–3–2             | 72.4 ± 0.9          | 39.7 ± 2.5 | 89.7 ± 1.5 | 92.9 ± 1.5      |
| 3–3–3             | 72.4 ± 1.8          | 40.4 ± 1.3 | 90.9 ± 0.8 | sym              |

The improvements in EC fabrics are observed for face masks with multiple layers. Aside from thickening the overall structure, the stacking of different EC fabrics could create a denser and more uniform distribution of pores, which would hinder airflow and particulate obstruction caused by overlapping layers of fibers. This study also found that the stacking combination of EC fabrics also led to an increase in $\Delta P$, up to between 67.5 and 253 Pa/cm². The overall performance is improved when the EC fabrics are stacked together in threes.

3.3. Prediction via machine learning

Using the physical characteristics of EC fabrics (Table 1) as features and $\Delta P$, PFE, and BFE (Table 3) as prediction targets, we trained Lasso and XGBoost models to predict $\Delta P$, PFE, and BFE of three new mask configurations. The results are compared with experimental values in Table 4; the root-mean-square errors (RMSEs) are shown in Table 5.

Table 4
Predicted and experimental values of $\Delta P$, PFE, and BFE with Lasso and XGBoost.

| Sample combination | $\Delta P$ (Pa/cm²) | PFE (%) | BFE (%) |
|-------------------|---------------------|---------|---------|
|                   | Experimental | Lasso    | XGBoost | Experimental | Lasso    | XGBoost | Experimental | Lasso    | XGBoost | Experimental | Lasso    | XGBoost |
| 1–2–2             | 142.1 ± 5.5 | 137.0 | 135.0  | 30.4 ± 0.7 | 42.0 | 35.4  | 94.8 ± 0.1 | 94.9 | 95.9  |
| 2–2–2             | 72.7 ± 3.5  | 67.5  | 67.0   | 28.6 ± 1.2 | 34.7 | 30.0  | 93.3 ± 1.4 | 93.9 | 94.2  |
| 2–3–1             | 147.4 ± 1.8 | 162.0 | 163.1  | 37.7 ± 0.3 | 35.9 | 40.0  | 94.8 ± 0.4 | 95.3 | 95.1  |

The predictions of BFE, one of the most important performance metrics, are consistently reliable with both Lasso and XGBoost: there is only a single prediction that is off by about 0.2% from the experimental value (combination 2–3–1, Table 4). On the other hand, the predictions of $\Delta P$ and PFE have a relatively larger deviation from the experimental value with larger RMSEs in general, though most of the predictions are still considered within the acceptable range given intrinsic experimental uncertainty and variability.

Overall, there is no significant difference in using either Lasso or XGBoost in predicting $\Delta P$, PFE, and BFE. The accuracy of predictions is mostly determined by the amount of training data available and the consistency of the training data, subjected to intrinsic experimental uncertainties.

3.4. Feature importance

To understand what are the key factors that are used by machine-learning models to make predictions, we analyzed the feature importance for each of the models. Especially, the feature importance for $\Delta P$, PFE, and BFE is calculated, and the top six features for each prediction are shown in Fig. 2 below. For Lasso, the weight of each feature is ranked and plotted; for XGBoost, the feature scores (f-score) are computed based on the information gain that each feature contributes after tree splits. It can be observed that the top features used by both machine-learning models are similar and are consistent with our expectations based on the physical knowledge of the materials. For example, the thickness is most influential to $\Delta P$, whereas PFE is mostly determined by pore sizes.

3.5. Correlations among $\Delta P$, PFE, and BFE

The correlation matrix of $\Delta P$, PFE, and BFE is also computed based on the experimental measurements on both the training and testing samples (Fig. 3). While $\Delta P$ is positively correlated with both PFE and BFE as expected, the correlation between PFE and BFE is rather weak at 0.20. This is consistent with the fact that bacterial filtration has a different capturing mechanism from particle
filtration due to size differences. The correlations between different features are also included in Fig. S6.

4. Discussion

4.1. Fabric properties vs. face mask efficacy

The general principle for basic aerosol filtration using physical barriers, such as face masks, is dependent on five mechanisms: gravity sedimentation, inertial impaction, interception, diffusion, and electrostatic attraction [11,20]. Gravitational settling, impaction, and interception are dominant mechanisms for larger size particles (>0.5 μm), while diffusion and electrostatic attraction prevail for smaller size particles (<0.2 μm) [20]. Based on these principles, we postulated a few correlations between the characteristics of EC fabric and the efficacy of protective face masks.

PFE, one important parameter of face mask design, was studied and assessed. Packing density is a determinant of PFE performance. With a more densely packed fiber network, there will be more surface area for droplet and aerosol deposition, enabling a higher capture efficiency [21]. The greatest PFE result belongs to EC100 (22.8%) in comparison to EC300 (17.6%) and EC200 (9.9%). Drewnick et al. also mentioned in their study that thread count is not proportional to filtration efficiency, because a larger thread count indicates thinner threads, which decrease material thickness [22], therefore resulting in poorer filtration efficiency. This was evident among the EC fabrics as EC100 (0.182 mm) is substantially thicker than EC200 (0.128 mm) and EC300 (0.120 mm). Taking into account the small particle size of 0.1 μm used in PFE testing, the dominant mechanisms active under this condition are diffusion and electrostatic attraction. The PFE outcome can be explained with the mean pore size obtained from porometry analysis. The diffusion capturing mechanism is effective for trapping particles of <0.2 μm [20]. Ascribing smaller mean pore size to more closely packed fibers, the probability of particles in random Brownian motion colliding with filter fibers is higher. Hence, it can be deduced that EC fabrics with smaller pore sizes result in a longer particle residence time, which improves PFE. The PFE results obtained support the deduction, with EC100 fabric of smallest mean pore size at 6.8 μm having the highest PFE outcome at 22.8%.

BFE, the most crucial parameter for protective face masks, was evaluated. Similarly, mean pore size has a direct impact on BFE. In consideration of the size of Staphylococcus aureus (3 μm), the likely predominant filtration methods are gravity sedimentation and inertial impaction [20]. The mean pore size of EC100 is the smallest at 6.8 μm and has the highest BFE outcome at 84.2%. A comparable finding was reported by Leanas and Jones whereby they found that the mask that has the best BFE also has the smallest mean pore size [23]. Although pore sizes for EC200 (14.9 μm) and EC300 (11.5 μm) are slightly different, BFE for the two fabrics are identical in the range of 48.3–50.6%. This could be attributed to similar mean pore sizes between the two; hence, there are no apparent effects on BFE results. The difference between the BFE of EC100 and that of the other two fabrics is mostly likely ascribed to a dense fiber network interference. Bacterial particles are then more prone to being entrapped by gravitational forces and higher inertia, making them unable to flow around the fibers.

Fig. 5. Feature importance in Lasso (red, top row) and XGBoost (blue, bottom row) models for ΔP (left panel), PFE (middle panel), and BFE (right panel). Thickness: fabric thickness, SA: specific surface area, fiberd: fiber strand diameter, porometer: mean pore size (pore_poro), pore: surface pore size (pore_BET), and zetap: zeta potential. The number at the end of a feature name indicates the layer number.
4.2. Mask performance prediction via machine learning

As shown in Table 4, BFE was predicted with the best accuracy, with the test RMSE comparable to the training RMSE, around 0.5% of the true values of BFE being predicted. For both the predicted results and the RMSE, there is no significant difference between Lasso and XGBoost models. Meanwhile, PFE was almost all predicted within the error range, but the average RMSE for testing is relatively large. In particular, the test RMSE was around 18% of the true values of PFE when Lasso was used. In this case, XGBoost generated a smaller RMSE. Despite the use of regularization and ensemble methods, there is large overfitting for ΔP and PFE, as seen from the test RMSEs being about ten times larger than the training RMSEs (Table 5).

Two main factors that have likely contributed to the less-than-ideal performance in predicting ΔP and PFE are a lack of training samples and the relatively larger intrinsic experimental variability and uncertainty. As seen from Table 3, the experimental uncertainties of triple-layered mask ΔP and PFE used for training are at least a few times larger than those of BFE. In addition, the variation in ΔP and PFE values are also larger than that in BFE. With these limitations, XGBoost was more robust than Lasso, and other regularized linear regression models. However, if only the size of the training set could be increased, Lasso and XGBoost are likely to have comparable performance.

One unique characteristic of this study is that a large number of features (21) were included, as the single-layer material properties that most directly influence a multilayer mask were unknown a priori. However, the coefficient w in Lasso and f-score in XGBoost provide information on the relative importance of features in predicting ΔP, PFE, and BFE (Fig. 2). For ΔP, single-layer thicknesses are more important. For BFE, both models suggest that single-layer thicknesses and fiber diameters carry far more weight than other features. However, for PFE, the pore sizes measured with either BET or the porometer are the determining factors. This contrast suggests that the capturing mechanism for various aerosol sizes is influenced by different physical characteristics.

Taking ΔP, PFE, and BFE into consideration, the thickness and the pore size of each mask layer are the most important fabric properties. In designing other types of multilayer masks in the future, one is likely to benefit the most by prioritizing the experimental measurement of thickness and pore size and using them as features in machine learning. In general, filtration efficiency is better with thicker fabrics and smaller pore sizes, considering the improved packing density and subsequent capture performance.

Recently, Lee et al. used deep learning to understand how fiber arrangement and the thickness of a single-layer N95 mask influence particulate filtration [24]. Findings from the study point toward better filtration efficiency with a fiber diameter of less than 1.8 μm and a denser fiber area. Neural network models were used in these studies, which are reliable but require a lot of training data. However, we have shown that given a limited data set, both Lasso and XGBoost are able to make good predictions of the properties of multilayer masks from only single-layer characteristics, with XGBoost having slightly better performance.

5. Conclusion

In this work, we have assessed the potential of a face mask made from EC. The fabric characteristics were analyzed, and the efficacy of the face mask was evaluated based on differential pressure, PFE, and bacterial filtration efficiency. Notably, the properties affecting packing density, such as fabric pore size, have an important role in determining face mask performance. A combination of three EC fabrics in different orders also brought about improved filtration effectiveness. The most outstanding combination is 1–3–1, which has the best PFE and BFE results at 45.4% and 98.1% respectively. Prediction outcomes from machine learning were accurate, especially for BFE. Despite limited training data, machine learning predicted all three properties with errors within experimental uncertainties, with BFE only differing 0.5% from the true values. Given the high frequency of viral mutations and the uncertainty associated with the efficacy of the vaccines, it is undeniable that masks will remain important in limiting the spread of the SARS-CoV-2 virus. ML algorithms are an inexpensive tool that complements experiments, speeding up the process of mask design.

Credit author statement

Yihao Leow, Jing K. Shi: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing — original draft, Writing — review & editing. Wei Liu, Xi Ping Ni, Pek Yin Michelle Yew: Methodology, Writing — review & editing. Songlin Liu, Zibiao Li: Methodology. Yang Xue, Dan Kai, Xian Jun Loh: Writing — review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors gratefully acknowledge the financial support from the Agency of Science, Technology and Research (A*STAR).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mtadv.2021.100178.

References

[1] M. Xie, Q. Chen, Insight into 2019 novel coronavirus – an updated interim review and lessons from SARS-CoV and MERS-CoV, Int. J. Infect. Dis. 94 (2020) 119–124.
[2] D. Daniel, M. Lin, I. Luhung, T. Lui, A. Sadovoy, X. Koh, A. Sing, T. Tran, S.C. Schuster, X. Jun Loh, O.S. Thet, C.K. Tan, Effective design of barrier enclosure to contain aerosol emissions from COVID-19 patients, Indoor Air 31 (5) (2021) 1639–1644.
[3] C.C. Ooi, A. Suwardi, Z.L.O. Yang, G. Xue, C.K.I. Tan, D. Daniel, H. Li, Z. Ge, F.Y. Leong, K. Marimuthu, O.T. Ng, S.B. Lim, P. Lim, W.S. Mak, W.C.D. Cheong, X.J. Loh, C.W. Kang, K.H. Lim, Risk assessment of airborne COVID-19 exposure in social settings, Phys. Fluids 33 (8) (2021), 087118.
[4] M.H. Chua, W.R. Cheng, S.S. Goh, J.H. Kong, B. Li, J.Y.C. Lim, L. Mao, S.S. Wang, K. Xue, L. Yang, E.Y. Ye, K.Y. Zhang, W.C.D. Cheong, B.H. Tan, Z.B. Li, B.H. Tan, X.J. Loh, Face masks in the new COVID-19 normal: materials, testing, and perspectives, Research 2020 (2020) 7286735.
[5] A. Suwardi, C.C. Ooi, D. Daniel, C.K. Tan, H.Y. Li, O.Y.Z. Liang, Y.K. Tang, J.Y. Chee, A. Sadovoy, S.Y. Jiang, S. Ramachandran, E.Y. Ye, C.W. Kang, W.C.D. Cheong, K.H. Lim, X.J. Loh, The efficacy of plant-based ionizers in removing aerosol for COVID-19 mitigation, Research 2021 (2021) 2173642.
[6] X. Jiang, Z. Li, D.J. Young, M. Liu, C. Wu, Y.L. Wu, X.J. Loh, Toward the prevention of coronavirus infection: what role can polymers play? Mater Today Adv 10 (2021) 100140.
[7] C.G. Wang, Z. Li, S. Liu, C.T. Ng, M. Marzuki, P.S. Jeslyn Wong, B. Tan, A. Lee, C.F. Hsi Lim, P. Bilani, Z. Fang, J.C. Chung Wong, Y.X. Setoh, Y.Y. Yang, C.H. Mun, S.Z. Fiona Phua, W.Q. Lim, L. Lin, A.R. Cook, H. Tanoto, L.C. Ng, A. Singhal, Y.W. Leong, X.J. Loh, N95 respirator decontamination: a study in reusability, Mater Today Adv 11 (2021) 100148.
[8] Q. Lin, J.Y.C. Lim, K. Xue, P.Y.M. Yew, C. Owh, P.L. Chee, X.J. Loh, Sanitizing agents for virus inactivation and disinfection, VIEW 1 (2) (2020) e16.
[9] R. Zhang, Y. Li, A.L. Zhang, Y. Wang, M.J. Molina, Identifying airborne transmission as the dominant route for the spread of COVID-19, Proc. Natl. Acad. Sci. U.S.A. 117 (26) (2020) 14857–14863.
A. Davies, K.-A. Thompson, K. Giri, G. Kafatos, J. Walker, A. Bennett, Testing the efficacy of homemade masks: would they protect in an influenza pandemic? Disaster Med. Public Health Prep. 7 (4) (2013) 413–418.

A. Konda, A. Prakash, G.A. Moss, M. Schmoldt, G.D. Grant, S. Guha, Aerosol filtration efficiency of common fabrics used in respiratory cloth masks, ACS Nano 14 (5) (2020) 6339–6347.

M. Tan, G. He, J. Nie, L. Zhang, L. Hu, Optimization of ultrafiltration membrane fabrication using backpropagation neural network and genetic algorithm, J. Taiwan Inst. Chem. Eng. 45 (1) (2014) 68–75.

L. Chu, J. Wang, T. Chu, Y. Qian, Z. Yu, D. Wu, Z. Zhang, Z. Jiang, J.O. Leckie, Modeling and optimizing the performance of PVC/PVB ultrafiltration membranes using supervised learning approaches, RSC Adv. 6 (33) (2016) 28038–28046.

R. Valaperta, M.R. Tejada, M. Frigerio, A. Moroni, E. Ciulla, S. Cioffi, P. Capelli, E. Costa, Staphylococcus aureus nosocomial infections: the role of a rapid and low-cost characterization for the establishment of a surveillance system, New Microbiol. 33 (3) (2010) 223–232.

F. Santosa, W.W. Symes, Linear inversion of band-limited reflection seismograms, SIAM J. Sci. Stat. Comput. 7 (4) (1986) 1307–1330.

R. Tibshirani, Regression shrinkage and selection via the Lasso, J. R. Stat. Soc. B 58 (1) (1996) 267–288.

T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, San Francisco, California, USA, 2016, pp. 785–794.

R.E. Schapire, The strength of weak learnability, Mach. Learn. 5 (2) (1990) 197–227.

How to interpret the output of XGBoost importance?. https://datascience.stackexchange.com/questions/12318/how-to-interpret-the-output-of-xgboost-importance, 2016. (Accessed 12 July 2021).

A. Tcharkhtchi, N. Abbasnezhad, M. Zarbini Seydani, N. Zirak, S. Farzaneh, M. Shirinbayan, An overview of filtration efficiency through the masks: mechanisms of the aerosols penetration, Bioact Mater 6 (1) (2021) 106–122.

S.-H. Huang, C.-W. Chen, Y.-M. Kuo, C.-Y. Lai, R. McKay, C.-C. Chen, Factors affecting filter penetration and quality factor of particulate respirators, Aerosol Air Qual. Res. 13 (1) (2013) 162–171.

F. Drewnick, J. Pikmann, F. Fachinger, L. Moormann, F. Sprang, S. Borrmann, Aerosol filtration efficiency of household materials for homemade face masks: influence of material properties, particle size, particle electrical charge, face velocity, and leaks, Aerosol. Sci. Technol. (2020) 1–17.

K. Leonas, C. Jones, The relationship of fabric properties and bacterial filtration efficiency for selected surgical face masks, J. Text. Apparel Technol. Manag. 3 (2003).

H.R. Lee, L. Liao, W. Xiao, A. Vailonis, A.J. Ricco, R. White, Y. Nishi, W. Chiu, S. Chu, Y. Cui, Three-dimensional analysis of particle distribution on filter layers inside N95 respirators by deep learning, Nano Lett. 21 (1) (2021) 651–657.